Maximum entropy random loopy graphs and imaginary replicas

Aguirre Lopez, Fabian

Awarding institution:
King's College London

Maximum Entropy Random Loopy Graphs and Imaginary Replicas

Fabián Aguirre López

Department of Mathematics

King's College London

A thesis submitted for the degree of

Doctor of Philosophy in Applied Mathematics

September 2020

Abstract

Random graph models have become extremely popular due to their relevance as models of complex networks. While the famous Erdös-Rényi model and the configuration model have been around for decades helping us gain enormous insight of the relevance of network structure for complex phenomena, these two models lack a fundamental element of real networks, the presence of short loops. Amazingly, there is no easy extension of traditional models to produce random loopy graphs that mimic real loopy networks. Only in the past decade have manageable loopy random graphs been properly developed. In this thesis we will present maximum entropy loopy models with a tuneable amount of short loops. Amazingly, even the simplest cases will present a rich phenomenology with nontrivial transitions. For the spectral calculations of these ensembles we introduce an extension of the traditional replica method that require certain replica limits where dimensions take imaginary values, contrary to the traditional $n \to 0$ case.

In the first part of the thesis we present a maximum entropy ensemble of random loopy graphs with a fixed degree sequence and a tuneable number of short loops. For the particular case of 2-regular graphs we present an exact solution of the model. It shows the range of tuneability of all loops of different lengths up to any finite value K. A transition from a connected to a disconnected phase is characterized, showing the nontrivial necessary scaling of the parameters with the system size. We extend the case of a triangle bias to a model with an arbitrary degree distribution. Through a simple yet powerful approximation of the generating function of the model we find a general theory applying to all degree distributions with well defined first and second moments. This ensemble undergoes a transition from a weak clustering regime where triangles do not interact, in the sense that they do not share edges, into a regime where triangles begin to cluster. We refer to this transition as the clustering transition. Eventually, triangles "break off" in the form of isolated cliques to maximize the triangle density. We refer to this last transition as the

shattering transition. We present accurate analytical estimates for the triangle density in the weak clustering regime. For the case of bounded degree distributions the scaling of the shattering transition with system size is also presented. We present the counter-intuitive result that every loop density will eventually fall in the shattered regime, no matter how small, if the graph is large enough. Overall we provide a complete picture of the behaviour of this ensemble validated by MCMC numerical sampling methods.

In the second part of the thesis we present the spectrally constrained maximum entropy ensemble. In this case we aim to bias the eigenvalue density at each point of the spectrum. For this we introduce a functional Lagrange parameter $\hat{\varrho}(\mu)$ that biases the spectral density at point μ. We develop an analytic theory based on the method of imaginary replicas in order to calculate the generating function of this model. Following insight from the first part, we develop the theory up to subleading terms in the system size, N, of $\mathcal{O}(1/N)$. We present a set of distributional equations that depend on the choice of model $\hat{\varrho}(\mu)$. We are able to recover the prediction for the loop density of the first part for the weak clustering regime, since the triangle bias is a particular case of this ensemble. With this theory we are also able to provide an analytic expression for the average eigenvalue density in the loopy regime. For the case of regular graphs an exact analytic solution is presented. For the case of an arbitrary degree distribution $p(k)$ the equations must be solved with numerical methods. A good agreement is found in all cases. The functional imaginary replica theory opens the door to other nontrivial ensembles beyond the triangle bias. These are explored easily for regular graphs thanks to the exact solution.

Acknowledgements

Finishing this PhD would not have been possible without the enormous support from my whole family in México, new and old friends, my very wise teachers, and many other people. I am deeply thankful with everyone that has helped me in many different situations on my long path towards the completion of this PhD for the past 30 years. The serious tone of a thesis does not allow me to show how there is a part of everyone in this text.

A special mention to all my friends, in London, Mexico, and everywhere in the world. Your friendship was a fundamental part of this PhD, as you all have contributed with support, ideas, and with the good times that have made this possible.

I would like to thank my supervisor, Professor Ton Coolen, for all the encouragement and trust over the last years. He was always there for me with good advice and pushing me in the right direction. A very special thanks to the people that were there everyday with me at the office, Alexander Mozeika and Mansoor Sheikh. This research project would not be what it is without your invaluable input, our endless discussions about any subject always pointed me in the right direction. Thanks to Dr Paolo Barucca who encouraged me to start this PhD. Special thanks to Dr Pierpaolo Vivo and Dr Alessia Annibale, who guided me and gave me the opportunity to teach their students. Thanks to everyone else at King's College London that have helped me since I arrived to the MSc program.

To Carlota, you have been with me since the very beginning of this adventure, thank you for all the immense loving support. I made it all this way thanks to you.

Above all, I would like to thank my mom, my dad, and my brother. You have supported me since forever and keep doing it tirelessly to this day. I dedicate this achievement to you. Thank you for being my family.

Lastly, I would like to thank the Consejo Nacional de Ciencia y Tecnología (CONACYT) from México for the scholarship that funded this PhD. Special thanks to all the Mexican workers that make these scholarships possible.

List of publications

1. **López, F. A.**, Barucca, P., Fekom, M., & Coolen, A. C. C. (2018). Exactly solvable random graph ensemble with extensively many short cycles. Journal of Physics A: Mathematical and Theoretical, 51(8), 085101.

   Material used in Chapter 2.

2. **López, F. A.**, & Coolen, A. C. C. (2020). Imaginary replica analysis of loopy regular random graphs. Journal of Physics A: Mathematical and Theoretical, 53(6), 065002.

   Material used in Chapter 5.

3. **Lopez, F. A.**, & Coolen, A. C. C. (2020). Transitions in loopy random graphs with fixed degrees and arbitrary degree distributions. arXiv preprint arXiv:2008.11002. Submitted to Journal of Physics: Complexity.

   Material used in Chapter 3.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Networks have proven to be fundamental in modelling many real world phenomena. If we go to the beginning of graph theory, it was introduced in 1741 to solve the problem of the seven bridges Königsberg, [1]. Euler noted that stripping all unnecessary detail of reality to solve this problem, the only thing left was a set of 4 nodes and 7 links between them, we call these type of abstract objects graphs now. Nowadays, in network science it is typical to work with far bigger graphs. But the objective is the same, take away a great deal of detail and filter down your problem to nodes and links between them. With the advent of powerful computers accessible to almost all researchers, it is normal for network scientists to work with real networks of nodes ranging from thousands to millions on a daily basis.

The field of network science and graph theory has proliferated so much precisely because there are in fact a lot of cases when a problem can be solved, or at least better understood, just by looking at the network structure alone. Just as in the case of the Köngisberg problem, for many problems we can strip all unnecessary detail and work only with nodes and links. At this point it is a good idea to mention that we will almost always speak of networks and graphs indistinctly. There is a tendency to refer to graphs as the abstract mathematical objects and to networks as the real observed objects, but we will not be strict about that difference. Mainly because we will not deal with any concrete "real world" problem in this thesis, but it is very important to keep this difference in mind. Especially because the use of a network theory approach to understand a real world problem is something that needs to be validated *a posteriori* and it should not be assumed to be sufficient *a priori*, as many of the researchers in the field of network science seem to believe. The fast and widespread use of network science tools is itself an important motivation of this thesis. It is fundamental to understand theoretically the tools that may

soon be used on a daily basis, otherwise we are at risk of blindly using tools that are only adequate in appearance.

When studying any phenomenon occurring on a graph one wants to understand how certain properties of the graph affect the process occurring. In some cases it may be possible to "solve" the problem for a particular instance of a graph, that is taking the whole information of the set of nodes and edges and produce a result either analytically or numerically. Nevertheless, from the point of view of theory, one might be interested in understanding how a particular property of the graph affects the processes of which it is the infrastructure. For this it is easier to imagine a family of graphs defined by a certain number of particular properties and see how the model behaves on a member of this family. For this we introduce the notion of *random graphs*.

By random graphs we refer to a family of graphs where each instance might occur with a certain probability. If the properties of these graphs can be tuned then it is possible to explore in a controlled way how certain properties affect models on networks by studying a very large amount of different graphs with the same properties.

1.1 Random graphs

1.1.1 First definitions

Formally a graph $G = (V, E)$ is an ordered pair of two sets. A set of N nodes, $V = \{1, \ldots, N\}$, and a set of edges between nodes, $E \subseteq V \times V = \{(i, j)|i, j \in V\}$. We will speak of edges or links indistinctly. While this is the formal definition we will work with graphs using the language of linear algebra, that is through the use of adjacency matrices. They will be denoted by $\mathbf{A}$ and are defined in the following way, $\mathbf{A} \in \{0, 1\}^{N \times N}$,

$$A_{ij} = \begin{cases} 1 & \text{if } (i, j) \in E \\ 0 & \text{if } (i, j) \notin E \end{cases} . \tag{1.1}$$

Therefore we can work with matrices instead of sets. Throughout the whole thesis we will only focus on a specific type of graphs called simple. This means that they do not contain any self connections, $(i, i) \notin E$, and that they are undirected, meaning that all links exist in both directions, $(i, j) \in E \Rightarrow (j, i) \in E$. In terms of adjacency matrices, it will mean that all matrices will be symmetric, $A_{ij} = A_{ji}$ and all diagonal elements are zero, $A_{ii} = 0$. The degree of a node is defined as the number of incident edges on it, $k_i(\mathbf{A}) = \sum_j A_{ij}$.

Needless to say, graphs are a very interesting combinatorial object themselves. As we will see, their study reveals fascinating theories connected to other branches of mathematics. Therefore, studying graphs *per se* is an important and interesting scientific endeavour. As a matter of fact most of the first part of the thesis uses heavily rigorous results from traditional random graph theory developed by abstract mathematicians.

To define an ensemble of random graphs we only need to define a probability distribution over the space of $0-1$ symmetric matrices with no diagonal elements. We will denote them by $p(\mathbf{A})$. Averages over the graph ensemble will be denoted by $\langle f(\mathbf{A})\rangle_{p(\mathbf{A})} = \sum_{\mathbf{A}} p(\mathbf{A}) f(\mathbf{A})$ or just simply $\langle f(\mathbf{A})\rangle$ if $p(\mathbf{A})$ can be inferred from the context.

One might wonder, why is it useful to define these ensembles? What can we learn about real networks from them? As scientists it is natural to ask a very simple question regarding the networks we have observed, what is typical and what is atypical? This is common for anyone with a statistical training as it is common to speak of null models in hypothesis testing, [2]. Random graphs are very good candidates to play the role of null models of real networks.

The simplest random graph ensembles focus only on the number of links and connectivity of each node. The simplest ensemble was introduced by Solomonoff and Rappoport in [3] and later reintroduced by Paul Erdös and Alfréd Rényi in [4], it is defined by making all edges statistically independent and identically distributed. It is famously referred as the Erdös-Rényi model (ER), defined in the following way.

ER Model

$$p_{ER}(\mathbf{A}) = \prod_{i<j} \left[\frac{c}{N} \delta_{A_{ij},1} + \left(1 - \frac{c}{N}\right) \delta_{A_{ij},0} \right]. \tag{1.2}$$

Where we have scaled with N the probability of each edge so that the expected number of edges is $\mathcal{O}(N)$. In this model the number of links is variable. Alternatively one can define an ensemble where all the degrees are fixed. It was introduced by Bollobas in [5], it is known as the configuration model (CM).

Configuration model

$$\begin{aligned} p_{CM}(\mathbf{A}) &= \frac{1}{\mathcal{N}_{\mathbf{k}}} \prod_{i=1}^{N} \delta_{k_i, \sum_j A_{ij}} \\ \mathcal{N}_{\mathbf{k}} &= \sum_{\mathbf{A}} \prod_{i=1}^{N} \delta_{k_i, \sum_j A_{ij}}. \end{aligned} \tag{1.3}$$

Here $\mathcal{N}_{\mathbf{k}}$ corresponds to the total number of graphs that have exactly the specified degree sequence $\mathbf{k} = (k_i)_{i=1,\dots,N}$. We will typically consider that the set of k_i's are sampled from a target degree distribution, $p(k)$. By the law of large numbers we know that for large graphs the empirical distribution will tend to the original one,

$$p(k|\mathbf{A}) = \frac{1}{N}\sum_{i=1}\delta_{k,k_i(\mathbf{A})} \xrightarrow{N\to\infty} p(k). \tag{1.4}$$

As simple as these ensembles may look, they have been object of intense study over the years. Their popularity is due mainly because models of interacting systems defined on their typical graph instances can often be solved analytically, and because sampling graphs from these ensembles is easy. Both properties derive mainly from the fact that ER and CM graphs are typically locally tree like (no short loops), which enables the application of many relatively simple mathematical and numerical approaches. In fact, nearly all mathematical and computational techniques currently available for analysing processes on large graphs (cavity methods, belief propagation and other message passing algorithms, generating functional analysis, conventional replica methods, etc) rely explicitly or implicitly on being able to neglect the presence of short loops, or on being able to treat such loops as perturbations of a fundamentally tree-like architecture.

Ironically it is precisely this property that makes the problems solvable the same property that makes them unrealistic. Real networks typically contain many short loops, [6, 7]. It also means that these models are useless as null models in the sense that every observed real network will be extremely atypical. Therefore we are required to add more properties to the graphs in a rational way, particularly loops. We deal with this in the following section.

1.1.2 Maximum entropy ensembles

The following step with random graph models was to introduce ensembles that included other observables, $\{f_\ell(\mathbf{A})\}_{\ell=1,\dots,K}$. We introduce the notion of maximum entropy (ME) random graph ensemble. If we want a random graph ensemble with the following expected averages,

$$\langle f_\ell(\mathbf{A})\rangle_{p(\mathbf{A})} = f_\ell^\star \qquad \text{for } \ell = 1,\dots,K. \tag{1.5}$$

We define it to be such that it maximizes the Shannon Entropy [8],

$$S[p] = -\sum_{\mathbf{A}} p(\mathbf{A}) \log p(\mathbf{A}). \tag{1.6}$$

In this way we guarantee that among all possible ensembles that satisfy (1.5), we are choosing the *fairest possible*. That is, the distribution of graphs is not biased by any other property besides the desired ones.

Solving the optimization problem over all possible distributions $p(\mathbf{A})$ we get an exponential family,

$$p(\mathbf{A}) = \frac{1}{Z(\boldsymbol{\alpha})} \exp\left(\sum_{\ell=1}^{K} \alpha_\ell f_\ell(\mathbf{A})\right). \tag{1.7}$$

Each α_ℓ should be chosen so that (1.5) is satisfied. If we choose the desired observable to be the total number of edges we recover the ER model, $f(\mathbf{A}) = \frac{1}{2}\sum_{ij}^{N} A_{ij}$. This means that the ER model is a maximum entropy ensemble.

Note that we can also impose hard constraints. This implies we want *all* the graphs to satisfy a certain set of properties. Therefore, we need to restrict the support of $p(\mathbf{A})$ to the subset of graphs that satisfied the desired constraints.

$$\forall_{m\in\{1,\dots,M\}} \;\; \forall_{\mathbf{A}\in \mathrm{supp}(p)} \qquad \Omega_m(\mathbf{A}) = \Omega_m^\star. \tag{1.8}$$

In this case the ME solution is simply the uniform distribution over all the graphs that satisfy these properties, this can be easily achieved by the use of delta distributions,

$$p(\mathbf{A}) = \frac{\prod_{m=1}^{M} \delta_{\Omega_m(\mathbf{A}),\Omega_m^\star}}{\sum_{\mathbf{A}'} \prod_{m=1}^{M} \delta_{\Omega_m(\mathbf{A}'),\Omega_m^\star}}. \tag{1.9}$$

If we choose the family of observables to be the degrees of all nodes then we recover the configuration model, $\Omega_i(\mathbf{A}) = k_i(\mathbf{A})$.

Finally, we can define a combination of soft and hard constraints giving rise to a mixed model. In this case the maximization of $S[p]$ has to be done only over the distributions that satisfy the hard constraint,

$$p(\mathbf{A}) = \frac{1}{Z(\boldsymbol{\alpha})} \exp\left(\sum_{\ell=1}^{K} \alpha_\ell f_\ell(\mathbf{A})\right) \prod_{m=1}^{M} \delta_{\Omega_m(\mathbf{A}),\Omega_m^\star}, \tag{1.10}$$

where the α_ℓ's need to be chosen in order to solve

$$f_\ell^\star = \sum_{\mathbf{A}} \frac{1}{Z(\boldsymbol{\alpha})} f_\ell(\mathbf{A}) \exp\left(\sum_{p=1}^{K} \alpha_p f_p(\mathbf{A})\right) \prod_{m=1}^{M} \delta_{\Omega_m(\mathbf{A}),\Omega_m^\star}. \tag{1.11}$$

In practice we will give the following interpretation to this ensemble. Consider the hard constraints to be a reference ensemble,

$$p_0(\mathbf{A}) = \frac{1}{Z_0} \prod_{m=1}^{M} \delta_{\Omega_m(\mathbf{A}),\Omega_m^\star}, \tag{1.12}$$

and then we can regard the set of α_ℓ's as *biases* with respect to the observable $f_\ell(\mathbf{A})$. This interpretation is useful as it is actually difficult to solve the set of equations (1.11). It is more common to simply study how the ensemble behaves as the parameters α_ℓ are varied.

In order to study ensembles like (1.10) it is useful to have a way to sample from them numerically. Luckily it has been found that through the Markov Chain Monte Carlo method (MCMC) it is possible to sample from many ensembles of that type. While this thesis will not go to deep into those methods, we refer the reader to [6, 9] for the main references of the sampling methods used in this book. A very brief recap is included in Appendix A.

1.1.3 Loopy random graphs

A random graph ensemble of loopy graphs is the natural improvement over the ER or the CM. The simplest and very intuitive idea is to simply add the number of triangles as an additional observable. This was done by Strauss in [10].

Strauss Model

$$p(A) = \frac{1}{Z(\lambda,\alpha)} \mathrm{e}^{\lambda L(\mathbf{A}) + \alpha T(\mathbf{A})}. \tag{1.13}$$

Where $L(\mathbf{A}) = \frac{1}{2}\sum_{ij} A_{ij}$ equals the number of links in the graph and $T(\mathbf{A}) = \frac{1}{6}\mathrm{Tr}(\mathbf{A}^3)$, which is exactly the number of triangles in the graph. As intuitive and simple as this model may be, it turns out it is of very little use. Strauss immediately noticed the ensemble condensed very quickly into dense graphs. And even though it has been proved that there is a small range of tuneability, [11], it is overall very hard to use. It quickly loses any

resemblance to real networks. There is a long history of studies aiming to understand this transition, [11–14], which was found to be a quite general feature of this type of ensemble with soft constraints.

There are two alternatives to solve the problem of the Strauss model. The first one is changing the ensemble into one with a hard constraint both on the number of single edges and the number of triangles around each node. This is the so called Newman-Miller model [15, 16]. Defining $s_i(\mathbf{A})$ as the number of single links around node i and $t_i(\mathbf{A})$ the number of triangles around node i, the model can be written as,

Newman-Miller Model

$$p(\mathbf{A}) = \frac{1}{Z}\prod_{i=1}^{N}\delta_{s_i,s_i(\mathbf{A})}\delta_{t_i,t_i(\mathbf{A})}. \tag{1.14}$$

Where the sequences $\{s_i\}$ and $\{t_i\}$ have to be sampled from given distributions $p(s)$ and $p(t)$. There have been important advances with this model both numerically and theoretically, studied before in [17–22] . Nevertheless, with this model it is difficult to keep the target degree distribution and the target total number of triangles under control, [23].

The other option is to keep the hard constraint on degrees but to add a soft constraint on the number of triangles.

$$p(\mathbf{A}) = \frac{1}{Z(\alpha)}\mathrm{e}^{\alpha\mathrm{Tr}(\mathbf{A}^3)}\prod_{i=1}^{N}\delta_{k_i,\sum_j A_{ij}}. \tag{1.15}$$

This is the ensemble we will study thoroughly in this thesis. As we will see, it will avoid the problems of the Strauss model but that does not mean it is perfectly controllable. As we shall see it has its own nontrivial transitions. The graphs change dramatically from a phase where triangles are not sharing edges or nodes into a clustered regime where triangles are packed together. This model was originally studied numerically in [24]. However, the sampling algorithm used in that case was incorrect as the acceptance probabilities chosen did not ensure convergence to the target distribution. This was pointed out in [6, 9], nontrivial acceptance probabilities are needed, as explained in Appendix A. Nevertheless, the overall phenomenology presented in [24] coincides with our results. A similar version but for dense Poissonian graphs, has also been studied. It also showed a similar phenomenology, [25–30]

1.2 The statistical physics approach and spectral methods

Th purpose of this thesis is mainly to advance the understanding the theoretical knowledge of loopy random graph ensembles. While we will rely on numerical sampling through MCMC, the main results presented here will be analytical. This is typically an extremely difficult task for loopy models. Rather than pursuing a rigorous approach we will make use of tricks from statistical physics to extract very accurate approximations that will describe analytically the models presented here.

We actually mean two different things when speaking of the *statistical physics approach*. The first refers to exploiting the fact that ME ensembles are of an exponential form, just like the Boltzmann distribution. This allows to take advantage of the intuition of statistical physics and of its jargon.

The main trick comes from the fact that in order to work with exponential models of this kind, (1.10), it is only necessary to calculate the *generating function* of the model. We will denote it by $\phi(\boldsymbol{\alpha})$ and define it as

$$\phi(\boldsymbol{\alpha}) = \frac{1}{N} \log \sum_{\mathbf{A}} \exp\left(\sum_{\ell=1}^{K} \alpha_\ell f_\ell(\mathbf{A}) \right) \prod_{m=1}^{M} \delta_{\Omega_m(\mathbf{A}),\Omega_m^\star}. \tag{1.16}$$

The expected value of all the observables that define the ensemble can be found by taking derivatives w.r.t. the corresponding Lagrange parameter, α_ℓ,

$$\frac{\partial}{\partial \alpha_\ell} \phi(\boldsymbol{\alpha}) = \frac{1}{N} \left\langle f_\ell(\mathbf{A}) \right\rangle. \tag{1.17}$$

A useful strategy is to insert δ's for the desired quantities,

$$\begin{aligned} \phi(\alpha) &= \frac{1}{N} \log \sum_{\boldsymbol{f}} \mathrm{e}^{\sum_\ell \alpha_\ell f_\ell} \mathcal{N}(\boldsymbol{f}) \\ \mathcal{N}(\boldsymbol{f}) &= \sum_{\mathbf{A}} \prod_{\ell=1}^{K} \delta_{f_\ell, f_\ell(\mathbf{A})} \prod_{m=1}^{M} \delta_{\Omega_m(\mathbf{A}),\Omega_m^\star}. \end{aligned} \tag{1.18}$$

As simple as this manipulation can look, it turns out it is extremely powerful. The change of focus from the generating function to the counting of states, $\phi(\alpha) \to \mathcal{N}(\boldsymbol{f})$, is extremely helpful when trying to do calculations analytically. In our case we will show its power in Chapters 2 and 3 to characterize very accurately the behaviour of the ensemble even beyond the $\mathcal{O}(1/N)$ regime, by using this trick in conjunction with known results from

the theory of uniform random graph models. It also plays a crucial role in the complicated calculations of the second part of the thesis, as it allows us to transform integrals into a form suitable for the saddle point approximation, [31].

The second application of statistical physics will be described in the following section where we actually introduce new variables to map our original problem into a new setting amenable to a statistical physics type calculation. But first we need to introduce notions of spectral graph theory.

1.2.1 Brief introduction to spectral graph theory

Spectral graph theory deals with studying the properties of the eigenvalues of the adjacency matrices of graphs. For our purposes, the main relationship to be used is that for simple graphs the traces of the adjacency matrix correspond to the number of closed paths of a given length in the graph. This can be easily proved if it is noted that the indicator function of a given path in a graph, $i \to j \to k \to \ell$, is precisely the product of the entries of the adjacency matrix for those links, $I(i \to j \to k \to \ell) = A_{ij}A_{jk}A_{k\ell}$.

$$\mathrm{Tr}(\mathbf{A}^\ell) = \#\text{ of closed paths of length } \ell. \tag{1.19}$$

Given that we are dealing with symmetric matrices, we know that the traces can be written in terms of the eigenvalues of $\mathbf{A}$,

$$\mathrm{Tr}(\mathbf{A}^\ell) = \sum_{i=1}^{N} [\mu_i(\mathbf{A})]^\ell, \tag{1.20}$$

where $\{\mu_i(\mathbf{A})\}_{i=1,\ldots,N}$ correspond to the ordered eigenvalues of $\mathbf{A}$. Since we will only be interested in quantities like (1.20) summed over all nodes, we can focus on calculating the spectral density or spectrum of the adjacency matrix only.

$$\varrho(\mu|\mathbf{A}) = \frac{1}{N}\sum_{i=1}^{N} \delta(\mu - \mu_i(\mathbf{A})). \tag{1.21}$$

In terms of this quantity the others like (1.20) are simply moments of this distribution.

$$\mathrm{Tr}(\mathbf{A}^\ell) = \sum_{i=1}^{N} [\mu_i(\mathbf{A})]^\ell = N \int \mathrm{d}\mu\, \mu^\ell \varrho(\mu|\mathbf{A}). \tag{1.22}$$

Nevertheless, note that $\varrho(\mu|\mathbf{A})$ is not exactly a function, since $\delta(x)$ is not really a

function. The Dirac delta, $\delta(x)$, can be formalized as a distribution or measure, but it is not a continuous function. Expression (1.21) is inconvenient since it requires the knowledge of each eigenvalue. Therefore we will hardly ever work directly with it. Instead we focus in the average spectral density.

$$\varrho(\mu) = \left\langle \frac{1}{N} \sum_{i=1}^{N} \delta(\mu - \mu_i(\mathbf{A})) \right\rangle. \tag{1.23}$$

This can become a continuous function. Nevertheless, it is still almost impossible to calculate for an arbitrary $p(\mathbf{A})$ for an arbitrary N. There are very nice examples where it is possible, but this is mostly for fully connected random matrices, [32]. The only other thing we can do is then to go to the so called asymptotic limit, $N \to \infty$. This is also a common strategy in statistical physics. It should be said that this strategy is only useful to compare with particular instances of very large graphs $\mathbf{A}$ if the model is self-averaging, in this case that means $Var[\varrho(\mu|\mathbf{A})] \xrightarrow{N\to\infty} 0$. This will be the case for the models studied here. Note that while we are not interested in infinite random graphs, we want to know a limiting behaviour that will be very close or almost indistinguishable from a finite instance. Actually for loop models it is very important not to take the limit $N \to \infty$ but rather to calculate the asymptotic value plus corrections of $\mathcal{O}(1/N)$.

We will not go into the full mathematical rigor of this field, but we will use many of the tools introduced for its study, see [32] for a nice introduction. Of these, the most important perhaps is the *Stieltjes transform*. Which is defined (in our convention) for an arbitrary measure $\varrho(\mu)$ in the following way,

$$S_\varrho(z) = \int \mathrm{d}\mu \, \varrho(\mu) \frac{1}{\mu - z}. \tag{1.24}$$

Note that while $\varrho(\mu|\mathbf{A})$, (1.21), is a sum of Dirac deltas and not a continuous function, its Stieltjes transform would not only be continuous but actually analytic! This will be extremely useful. We define the Stieltjes transform for the spectral density as

$$S_N(z|\mathbf{A}) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\mu_i(\mathbf{A}) - z} = \frac{1}{N} \mathrm{Tr}[(\mathbf{A} - z\mathbf{I})^{-1}]. \tag{1.25}$$

While it is still hard to calculate for an arbitrary matrix it is a much more manageable expression than (1.21) since it involves an explicit formula and it is analytic.

The Stieltjes transform can be inverted in the following way,

$$\varrho(\mu) = \lim_{\varepsilon\to 0} \frac{1}{\pi}\mathrm{Im}\, S_\varrho(\mu + \mathrm{i}\varepsilon). \tag{1.26}$$

This is useful because in order to calculate the average spectral density (1.23), we can do it by first calculating the Stieltjes transform, then doing the ensemble average, and then inverting the result with (1.26). Abbreviating $\mu_\varepsilon = \mu + \mathrm{i}\varepsilon$,

$$\varrho(\mu) = \lim_{\varepsilon\to 0} \frac{1}{\pi}\mathrm{Im}\left\langle \mathrm{Tr}\left[(\mathbf{A} - \mu_\varepsilon \mathbf{I})^{-1}\right]\right\rangle_{p(\mathbf{A})}. \tag{1.27}$$

This formula as it is combined with large N arguments is extremely useful to calculate asymptotic spectral densities. See [32] for examples. For the case of graphs, combined with the Schur complement formula, (1.25) is the basis for the cavity method for spectral densities, [33], valid for instances of locally tree-like graphs.

More recently (1.25) has been used as the starting point for spectral calculations of loopy graphs. A formalism of message passing with loops was shown in [22], allowing for spectral calculations of instances of the Newman-Miller model. Average spectral densities of the Newman-Miller model with a small amount of loops were calculated in [20] starting from (1.27). Spectral densities of regular trees of loops where calculated in [34] also starting from (1.25) in combination with the Schur complement formula. This approach turns out to be closely related to the replica formalism we develop in the thesis as we will mention in further chapters.

While this form (1.27) is useful, it still needs to be transformed to be used with the replica formalism in the context of statistical physics.

As a final commentary we point out that actually (1.25) can be used as a smooth approximation of the spectral density when the imaginary part of the argument is small.

$$\begin{aligned}\varrho_\varepsilon(\mu|\mathbf{A}) =& \frac{1}{\pi}\mathrm{Im}\, S(\mu_\varepsilon|\mathbf{A}) = \frac{1}{N}\sum_{i=1}^{N}\delta_\varepsilon(\mu - \mu_i(\mathbf{A}))\\ \delta_\varepsilon(\mu) =& \frac{1}{\pi}\frac{\varepsilon}{\mu^2+\varepsilon^2}\end{aligned} \tag{1.28}$$

Thanks to the Sokhotski–Plemelj formula, [35], this is actually a very good approximation for the spectral density, not only visually, but it is also useful to calculate integrals over it as long as the domain of integration is finite. In that way we avoid the divergent tails

of the Cauchy distributions δ_ε.

$$\int_a^b \mathrm{d}\mu f(\mu)\varrho(\mu) = \lim_{\varepsilon\to 0}\int_a^b \mathrm{d}\mu f(\mu)\varrho_\varepsilon(\mu) \tag{1.29}$$

This means we can approximate the integral on the LHS of (1.29) with any desired accuracy with an integral over $\varrho_\varepsilon(\mu)$, provided ε is small enough.

1.2.2 Replica formalism for spectral densities

In order to bring the spectral calculation into the realm of statistical physics it is only necessary to note that (1.25) can be converted into the following expression.

$$\begin{aligned}\varrho(\mu|\mathbf{A}) &= \lim_{\varepsilon\to 0}\frac{2}{\pi N}\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log Z(\mu_\varepsilon|\mathbf{A}),\\ Z(\mu_\varepsilon|\mathbf{A}) &= \int\prod_{i=1}^N \mathrm{d}\phi^i \exp\Big[-\frac{\mathrm{i}}{2}\sum_{ij}\phi^i(A_{ij}-\mu_\varepsilon\delta_{ij})\phi^j\Big].\end{aligned} \tag{1.30}$$

This is the famous Edwards-Jones (EJ) formula , [36]. We say it has been mapped into a statistical physics setting because the Gaussian integral $Z(\mu_\varepsilon|\mathbf{A})$ can be regarded as the partition function of N interacting soft spins $\{\phi^i\}_{i=1,\dots,N}$ with Hamiltonian given by $H(\boldsymbol{\phi}) = \frac{\mathrm{i}}{2}\sum_{ij}\phi^i(A_{ij}-\mu_\varepsilon\delta_{ij})\phi^j$. This is almost like actual problems studied in statistical physics, see [37]. The main difference is that the Hamiltonian is complex in this case, but the analogy is very useful. One important consequence of having a complex Hamiltonian is that the log in (1.30) is not uniquely defined as it is a complex logarithm. It is sometimes incorrectly claimed that if the principal branch is chosen, Log, then the formula is valid. However, a careful analysis reveals that spurious negative δ's appear when taking the derivative $\frac{\mathrm{d}}{\mathrm{d}\mu}$, and that they do not disappear even after taking the ensemble average. The appropriate way of fixing the branch of this complex logarithm to the correct value is by remembering that after taking the derivative $\frac{\mathrm{d}}{\mathrm{d}\mu}\log Z(\mu_\varepsilon|\mathbf{A})$ it should match (1.25). That means $\log Z(z|\mathbf{A})$ should be an *analytic function* and its derivative should be $\frac{\mathrm{d}}{\mathrm{d}z}\log Z(z|\mathbf{A}) = (1/2)S(z|\mathbf{A})$, while this may pass as a triviality this is actually an active choice of a branch, since finding $\frac{\mathrm{d}}{\mathrm{d}z}\log f(z) \neq f'(z)/f(z)$ is possible for complex numbers. The equality is only true on a certain branch. This will be further discussed in Chapter 4.

The most important use of (1.30) is actually to calculate average spectral densities, $\langle\varrho(\mu|\mathbf{A})\rangle$. This because we have changed the role of $\mathbf{A}$ in the calculation. Before we were looking at $\mathbf{A}$ as the variable of study, in the context of (1.30) $\mathbf{A}$ is actually a disorder

in the Hamiltonian and the variable is the vector $\boldsymbol{\phi}$. This is important as it is precisely this that will allow us to use the replica trick from statistical physics, [38]. According to (1.30), the average spectral density can be calculated in the following way,

$$\begin{aligned}\varrho(\mu) =& \lim_{\varepsilon\to 0}\frac{2}{\pi}\mathrm{Im}\frac{\mathrm{d}}{\mathrm{d}\mu}\phi(\mu_\varepsilon),\\ \phi(\mu_\varepsilon) =& \frac{1}{N}\left\langle \log Z(\mu_\varepsilon|\mathbf{A})\right\rangle .\end{aligned} \tag{1.31}$$

In this case $\phi(\mu_\varepsilon)$ corresponds to minus the average free energy density of our system of soft spins $\boldsymbol{\phi}$ with disorder $\mathbf{A}$. The replica trick, [38], was introduced precisely for this, to calculate correct asymptotic expressions for the free energy density averaged over the disorder. It is implemented in the following way,

$$\phi(\mu_\varepsilon) = \frac{1}{N}\left\langle \log Z(\mu_\varepsilon|\mathbf{A})\right\rangle = \lim_{n\to 0}\frac{1}{Nn}\log\left\langle Z(\mu_\varepsilon|\mathbf{A})^n\right\rangle . \tag{1.32}$$

The trick consists of first calculating the last expression for positive integers n, where Z^n can be written as a multiple integral and thus the average over $p(\mathbf{A})$ can be performed analytically.

$$\left\langle Z(\mu_\varepsilon|\mathbf{A})^n\right\rangle = \int\prod_{ia}\mathrm{d}\phi_a^i\,\mathrm{e}^{\frac{\mathrm{i}}{2}\mu_\varepsilon\sum_{ia}(\phi_a^i)^2}\left\langle \mathrm{e}^{-\frac{\mathrm{i}}{2}\sum_{ij}A_{ij}\sum_a\phi_a^i\phi_a^j}\right\rangle_{p(\mathbf{A})} . \tag{1.33}$$

We will discuss at length how these integrals are calculated in the second part of the thesis. They are usually solved by calculating an asymptotic result, that is,

$$\left\langle Z(\mu_\varepsilon|\mathbf{A})^N\right\rangle \sim \mathrm{e}^{Nf(\mu_\varepsilon|n)} \iff \lim_{N\to\infty}\frac{\left\langle Z(\mu_\varepsilon|\mathbf{A})^N\right\rangle}{\mathrm{e}^{NS(\mu_\varepsilon|n)}} = 1. \tag{1.34}$$

Which means we can calculate the asymptotic value for the average spectral density by substituting in (1.32).

$$\varrho(\mu) \sim \lim_{\varepsilon\to 0}\lim_{n\to 0}\frac{2}{Nn\pi}\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log\mathrm{e}^{Nf(\mu_\varepsilon|n)} = \lim_{\varepsilon\to 0}\lim_{n\to 0}\frac{2}{\pi n}\mathrm{Im}f'(\mu_\varepsilon|n). \tag{1.35}$$

Where we have assumed $\frac{\mathrm{d}}{\mathrm{d}\mu}\log\mathrm{e}^{Nf(\mu_\varepsilon|n)} = Nf'(\mu_\varepsilon|n)$, note that this also involves a choice of branch, which we discuss further in Chapter 4. Even though there are a few more technical details, overall the statistical physics approach offers a very clear path to calculating the desired observable $\varrho(\mu)$. It has been successfully used many times, in particular for

graphs in [39–41]. In our case we will actually take this procedure one step further and take the replica limit not to 0 but to an imaginary value. We leave the presentation of that formula to Chapter 4, which is devoted to present that whole procedure.

Calculating only for integer n and then assuming it has a continuous value has been contested many times, in particular for spectral densities in [42, 43], nevertheless there is no evidence of it giving wrong results for spectral densities. Even in the case of graphs it has been proven to be equivalent to other more rigorous methods like the super-symmetric one, [44]. It is important to remember that without a formal proof one has to be cautious and always consider the possibility of the replica trick giving the wrong result. But this is not our concern as we are trying to find expressions to test against numerical simulations and not aiming to write down proofs.

As a curious historical note, the replica trick was invented in a different form by Mark Kac, [45], for the problem of calculating the spectral density of modes of a disordered spring chain. The result obtained was actually a first version of the general equation more recently described in [39].

1.3 Outline of the thesis

The thesis can be summarized very quickly in the following three bullet points:

- In the first part we study and characterize a random graph ensemble with a tuneable number of triangles and fixed degree sequence.
- We calculate its spectral properties by introducing a functional formalism with imaginary replicas.
- On the way we show the functional formalism has the potential of being applied to more complicated ensembles than just a triangle bias.

The thesis is divided in two parts. The first part deals with studying the random graph ensembles with a triangle bias.

In Chapter 2 we present the simple possible nontrivial random loopy graph ensemble. It consists of random 2-regular graphs with a bias for short loops. While these graphs might seem too simple, they turn out to be extremely useful as they already show many similarities with more general models. The ensemble also possesses an exact asymptotic solution which is formally similar to the solution of the more complicated ensembles.

Chapter 3 deals with a general loopy graph ensemble. It consists of a mix between a hard constraint on all degrees, in the same way as the configuration model, plus a bias for triangles. The phenomenology turns out to be rich and interesting. The general behaviour observed is a transition from an initial regime of noninteracting triangles into a clustered regimes where triangles are packed together is observed. Accurate analytic estimates validated by MCMC simulations are shown.

In the second part of the thesis the imaginary replica functional formalism is presented.

In Chapter 4 the formalism is introduced in such a way that is valid for random matrices in general. There is a step by step derivation of the main calculations to show similarities and dissimilarities with the traditional replica method. As this formalism relies on many non rigorous assumptions the objective is to generate a clean list of all the procedure for further verification.

In Chapter 5 the imaginary replica functional formalism is applied to regular graphs. The results from previous chapters are reproduced and the average spectral density is calculated. The leading order and subleading orders in N, the system size, are successfully obtained analytically. The main advantage of the regular case is that it allows for exact solutions. The results are compared with extensive numerical simulations, including tests with more complicated models.

In Chapter 6 the results from the previous chapter are generalized to an arbitrary degree distribution. The fact that the degree distribution is not specified eliminates the possibility of an exact analytic solution. Nevertheless it is possible to extract some results that are successfully compared against numerical simulations.

The last chapter presents conclusions and outlooks.

Part I

Maximum Entropy Loopy Graphs - Direct Approach

Chapter 2

Exactly solvable loopy model

In this chapter we introduce the simplest possible loopy random graph model. It consists of biasing 2-regular graphs with respect to short loops. The simplicity of the possible loopy graph topologies in this case allows one to get an exact solution. Amazingly, the phenomenology of this model is similar to the more general model studied in the next chapter. The calculation presented here also served as inspiration for the one of the general case.

2.1 The model

We define a random graph ensemble over the set of undirected simple regular graphs of degree 2, which we denote by $\mathcal{G}_N$. Any graph in $\mathcal{G}_N$ is necessarily a set of disjoint closed loops. The probability assigned to each graph $\mathbf{A} \in \mathcal{G}_N$ is chosen proportional to the exponential of a weighted sum of the number of triangles, squares, pentagons, ... , K-loops present in $\mathbf{A}$. We refer to this as biasing with respect of the number of short loops. Thus

$$p(\mathbf{A}) = \frac{1}{Z_N(\boldsymbol{\beta})} \exp\left(\sum_{\ell=3}^{K} \ell \beta_\ell n_\ell(\mathbf{A})\right), \tag{2.1}$$

Here $n_\ell(\mathbf{A})$ denotes the number of length ℓ-loops, i.e. closed paths of length ℓ without backtracking and without over-counting, and $\boldsymbol{\beta} = (\beta_3, \ldots, \beta_K) \in \mathbb{R}^{K-2}$ is a vector of control parameters. Note that isolated nodes ($\ell = 1$) and dimers ($\ell = 2$) cannot occur due to the degree constraint. The factors ℓ in (2.1) are included in the exponent for later convenience. We are effectively biasing with respect to the total number of ℓ-loops starting

at a given node through the introduction of the field β_ℓ.

The partition function $Z_N(\boldsymbol{\beta})$ is given by

$$Z_N(\boldsymbol{\beta}) \;\; = \sum_{\mathbf{A}\in\mathcal{G}_N} \exp\left(\sum_{\ell=3}^{K} \ell\beta_\ell n_\ell(\mathbf{A})\right). \tag{2.2}$$

Expression (2.1) defines a maximum entropy random graph ensemble with respect to the $K-2$ observables $n_\ell(\mathbf{A})$, whose ensemble averages are controlled by varying the parameters $\boldsymbol{\beta}$. We choose K to be a fixed number for all values of N. This exponential form is a particular version of the one presented in equation (1.1) of [13]. It is an ensemble where we are interested in controlling the expected values of a finite number of graph observables.

The average fraction of the N nodes that will be found in an ℓ-loop is given by

$$m_\ell \;\; = \;\; \frac{\ell}{N}\langle n_\ell(\mathbf{A})\rangle. \tag{2.3}$$

where $\langle f(\mathbf{A})\rangle = \sum_{\mathbf{A}} p(\mathbf{A}) f(\mathbf{A})$. Following the statistical mechanics route, we define a generating function $\phi_N(\boldsymbol{\beta})$:

$$\phi_N(\boldsymbol{\beta}) \;\; = \;\; N^{-1}\log[Z_N(\boldsymbol{\beta})/N!]. \tag{2.4}$$

The main quantities of interest (2.3) for our graph ensemble (2.1) can be computed from (2.4) via

$$m_\ell \;\; = \;\; \partial\phi_N(\boldsymbol{\beta})/\partial\beta_\ell. \tag{2.5}$$

The generating function $\phi_N(\boldsymbol{\beta})$ is the free entropy, apart from a factor $N!$ which reflects (topologically irrelevant) node label permutations. Including this factor will ensure that the limit $\phi(\boldsymbol{\beta}) = \lim_{N\to\infty}\phi_N(\boldsymbol{\beta})$ exists.

2.2 Analytical solution

2.2.1 Summation over graphs

To evaluate the partition function (2.2) we need to perform a sum over graphs. Such sums are usually not analytically tractable, especially when the ensemble definition involves loops, as is the case in (2.1). Here we are able to perform the summation by rewriting it

as

$$Z_N(\boldsymbol{\beta}) = \sum_{\boldsymbol{n}} D(\boldsymbol{n}) \mathrm{e}^{\sum_{\ell=3}^{K} \ell \beta_\ell n_\ell}, \tag{2.6}$$

with $\boldsymbol{n} = (n_3, \ldots, n_N) \in \mathbb{N}^{N-2}$. This decomposition reflects the fact that, in the particular case of $\mathcal{G}_N$, we are fortunate that each graph has to be a collection of loops, and can therefore be identified fully by a sequence $\boldsymbol{n} = (n_3, \ldots, n_N)$ that specifies the number of loops of each possible length up to N, and a labelling of the nodes. The sum over graphs is then performed by summing over all possible sequences $\boldsymbol{n}$, keeping track of the multiplicity of each sequence via an associated density of states $D(\boldsymbol{n})$:

$$D(\boldsymbol{n}) \;=\; \sum_{\mathbf{A} \in \mathcal{G}_N} \prod_{\ell=3}^{N} \delta_{n_\ell, n_\ell(\mathbf{A})} = \frac{N!\; \delta_{N, \sum_{\ell=3}^{N} \ell n_\ell}}{\prod_{\ell=3}^{N} [(2\ell)^{n_\ell} n_\ell!]}. \tag{2.7}$$

Apart from the condition $N = \sum_{\ell=3}^{N} \ell n_\ell$, this density is proportional to $N!$ but corrected for over-counting due to the indistinguishability of different length-ℓ loops, giving a divisor $n_\ell!$, and due to the different ways one can number the nodes in each ℓ-loop without altering the graph (ℓ cyclic permutations, plus ℓ anti-cyclic permutations), giving a further divisor $(2\ell)^{n_\ell}$. Using the integral form of the Kronecker delta $\delta_{nm} = \int_{-\pi}^{\pi} (\mathrm{d}\omega/2\pi) \mathrm{e}^{\mathrm{i}\omega(n-m)}$, we can thus write the partition function as

$$\begin{aligned} Z_N(\boldsymbol{\beta}) \;&=\; \sum_{\boldsymbol{n}} \frac{N!}{\prod_{\ell=3}^{N} [(2\ell)^{n_\ell} n_\ell!]} \left(\prod_{\ell=3}^{K} \mathrm{e}^{\ell \beta_\ell n_\ell} \right) \int_{-\pi}^{\pi} \frac{\mathrm{d}\omega}{2\pi} \mathrm{e}^{\mathrm{i}\omega \left(N - \sum_{\ell=3}^{N} \ell n_\ell \right)} \\ &=\; \frac{N!}{2\pi} \int_{-\pi}^{\pi} \mathrm{d}\omega \; \mathrm{e}^{\mathrm{i}\omega N} \prod_{\ell=3}^{K} \left(\sum_{n_\ell \geq 0} \frac{\mathrm{e}^{(\beta_\ell - \mathrm{i}\omega)\ell n_\ell}}{(2\ell)^{n_\ell} n_\ell!} \right) \prod_{\ell=K+1}^{N} \left(\sum_{n_\ell \geq 0} \frac{\mathrm{e}^{-\mathrm{i}\omega \ell n_\ell}}{(2\ell)^{n_\ell} n_\ell!} \right) \\ &=\; \frac{N!}{2\pi} \int_{-\pi}^{\pi} \mathrm{d}\omega \; \exp\left(\mathrm{i}\omega N + \sum_{\ell=3}^{K} \frac{\mathrm{e}^{(\beta_\ell - \mathrm{i}\omega)\ell}}{2\ell} + \sum_{\ell=K+1}^{N} \frac{\mathrm{e}^{-\mathrm{i}\omega\ell}}{2\ell} \right). \end{aligned} \tag{2.8}$$

From this, in combination with (2.4), we infer that

$$\phi_N(\boldsymbol{\beta}) \;=\; \frac{1}{N} \log \int_{-\pi}^{\pi} \frac{\mathrm{d}\omega}{2\pi} \; \mathrm{e}^{N f_N(\omega, \boldsymbol{\beta})} \tag{2.9}$$

with

$$f_N(\omega, \boldsymbol{\beta}) \;=\; \mathrm{i}\omega + \sum_{\ell=3}^{K} \frac{\mathrm{e}^{(\beta_\ell - \mathrm{i}\omega)\ell}}{2\ell N} + \sum_{\ell=K+1}^{N} \frac{\mathrm{e}^{-\mathrm{i}\omega\ell}}{2\ell N}. \tag{2.10}$$

An exact expression for (2.9), valid for any finite N, would require to perform the integral in it. Instead, we proceed in the usual way as in statistical physics. We look at the thermodynamic limit, focusing then on $\phi(\boldsymbol{\beta}) = \lim_{N\to\infty}\phi_N(\boldsymbol{\beta})$. This will allow us to calculate the asymptotic expressions for (2.3), which should differ from the finite size values by $\mathcal{O}(1/N)$ corrections.

The limit $N \to \infty$ of (2.9) can now be obtained by evaluating the integral over ω in (2.8) via steepest descent:

$$\phi(\boldsymbol{\beta}) = \lim_{N\to\infty} \mathrm{extr}_\omega f_N(\omega, \boldsymbol{\beta}). \tag{2.11}$$

The extremum is found by solving $\partial f_N(\omega, \boldsymbol{\beta})/\partial\omega = 0$.

2.2.2 Scaling with N control parameters

We observe that for finite $\{\beta_\ell\}$ our model cannot exhibit nonzero loop densities m_ℓ in the infinite size limit, since the $\boldsymbol{\beta}$-dependent term in (2.10) vanishes for $N \to \infty$. We are therefore led to redefining the parameters $\boldsymbol{\beta}$ with a size dependent shift,

$$\beta_\ell \;=\; \tilde{\beta}_\ell + \frac{1}{\ell}\log(N), \tag{2.12}$$

where $\tilde{\beta}_\ell = \mathcal{O}(1)$. An intuitive explanation for this scaling is presented in section 2.3. We denote the vector of shifted $\mathcal{O}(1)$ control parameters by $\tilde{\boldsymbol{\beta}} = (\tilde{\beta}_3, \ldots, \tilde{\beta}_K)$, and we define $\phi_N(\boldsymbol{\beta}) = \varphi_N(\tilde{\boldsymbol{\beta}})$. This implies that for $N \to \infty$ we will have $m_\ell = \partial\varphi(\tilde{\boldsymbol{\beta}})/\partial\tilde{\beta}_\ell$, in which now

$$\varphi(\tilde{\boldsymbol{\beta}}) \;=\; \lim_{N\to\infty} \mathrm{extr}_\omega\Big\{\mathrm{i}\omega + \sum_{\ell=3}^{K} \frac{\mathrm{e}^{(\tilde{\beta}_\ell - \mathrm{i}\omega)\ell}}{2\ell} + \sum_{\ell=K+1}^{N} \frac{\mathrm{e}^{-\mathrm{i}\omega\ell}}{2\ell N}\Big\}. \tag{2.13}$$

Differentiation of this latter expression reveals that the value ω_N at the extremum is to be solved from

$$1 \;=\; \frac{1}{2}\sum_{\ell=3}^{K} \mathrm{e}^{(\tilde{\beta}_\ell - \mathrm{i}\omega_N)\ell} + \frac{1}{2N}\sum_{\ell=K+1}^{N} \mathrm{e}^{-\mathrm{i}\omega_N\ell}, \tag{2.14}$$

and that the asymptotic values of the observables m_ℓ are subsequently given by

$$m_\ell = \frac{1}{2}\mathrm{e}^{(\tilde{\beta}_\ell - \mathrm{i}\omega_N)\ell}. \tag{2.15}$$

This last identity, in combination with (2.14), prompts us to introduce $m_\infty = 1 - \sum_{\ell \leq K} m_\ell$, with $m_\ell \in [0,1]$, which gives the fraction of the nodes that are *not* in loops of length K or less. It is for $N \to \infty$ apparently given by

$$m_\infty \;=\; \lim_{N\to\infty} \frac{1}{2N} \sum_{\ell=K+1}^{N} \mathrm{e}^{-\mathrm{i}\omega_N \ell}. \tag{2.16}$$

It follows from (2.15), that the physical saddle point ω, after contour deformation, must be purely imaginary. We switch accordingly to the new variable $x = \mathrm{e}^{-\mathrm{i}\omega} \in \mathbb{R}_0^+$, in terms of which our equations become:

$$1 \;=\; \frac{1}{2} \sum_{\ell=3}^{K} x_N^\ell \mathrm{e}^{\ell\tilde{\beta}_\ell} + \frac{1}{2N} \sum_{\ell=K+1}^{N} x_N^\ell, \tag{2.17}$$

$$m_\ell \;=\; \lim_{N\to\infty} \frac{1}{2} x_N^\ell \mathrm{e}^{\ell\tilde{\beta}_\ell}, \tag{2.18}$$

$$m_\infty \;=\; \lim_{N\to\infty} \frac{1}{2N} \sum_{\ell=K+1}^{N} x_N^\ell. \tag{2.19}$$

2.2.3 Phase phenomenology of the ensemble

We will now demonstrate that the solutions to the coupled equations (2.17,2.18) give rise to two phases of our graph ensemble. A *disconnected* phase is characterized by the fact that all nodes are typically in loops of length K or less, so $m_\infty = 0$. A second phase, the *connected* phase, is characterized by finding a finite fraction of the nodes in longer loops, so here $m_\infty > 0$. The transition separating the phases is marked by bifurcation of $m_\infty > 0$ solutions.

If $\lim_{N\to\infty} x_N = x < 1$, the second term of (2.17) vanishes for $N \to \infty$, and we immediately obtain $m_\infty = 0$. Hence we are in the disconnected phase, and here the asymptotic observables m_ℓ are simply found by solving

$$m_\infty = 0: \qquad 1 = \frac{1}{2} \sum_{\ell=3}^{K} x^\ell \mathrm{e}^{\ell\tilde{\beta}_\ell}, \quad m_\ell = \frac{1}{2} x^\ell \mathrm{e}^{\ell\tilde{\beta}_\ell}. \tag{2.20}$$

The condition $x < 1$ for this solution to exist will be met for large values of $\{\tilde{\beta}_\ell\}$. Upon reducing the control parameters $\{\tilde{\beta}_\ell\}$, the value of x will increase, and a transition to the connected phase occurs exactly when $x = 1$. This happens at the critical manifold in the

$K-2$ dimensional parameter space, defined by validity of

$$\sum_{\ell=3}^{K} \mathrm{e}^{\ell\tilde{\beta}_\ell} = 2. \tag{2.21}$$

To confirm the equations of the connected phase, we need to investigate how the solution x_N of (2.17) scales with N as we approach $x = 1$. Substituting $x_N = 1 - \xi/N$, expanding (2.17) in N, and taking the limit $N \to \infty$ gives

$$m_\ell = \frac{1}{2}\mathrm{e}^{\ell\tilde{\beta}_\ell}, \qquad m_\infty = 1 - \frac{1}{2}\sum_{\ell=3}^{K} \mathrm{e}^{\ell\tilde{\beta}_\ell}, \tag{2.22}$$

and the link between ξ and m_∞ is $m_\infty = (1-\mathrm{e}^{-\xi})/2\xi$.

It turns out that all loops of finite length $L > K$ will always have vanishing density for $N \to \infty$. This can be seen simply by replacing $K \to L$ in the previous analysis, but with $\beta_\ell = 0$ for all $K < \ell \leq L$. The newly added control parameters with $\ell > K$ will give $\tilde{\beta}_\ell = -\ell^{-1}\log N$, and hence $m_\ell = \lim_{N\to\infty} \frac{1}{2}x^\ell \mathrm{e}^{\ell\tilde{\beta}_\ell} = \lim_{N\to\infty} \frac{1}{2}x^\ell/N = 0$, in both phases. We knew that in the disconnected phase all nodes will typically be in the controlled short loops of length K or less. We may now conclude that, in the connected phase, those nodes that are not in the controlled short loops (the fraction $m_\infty > 0$) will typically be found in loops of *diverging* length.

The ensemble's Shannon entropy [8] is given by

$$\begin{aligned} S_N &= -\sum_{\mathbf{A}\in\mathcal{G}_N} p(\mathbf{A})\log p(\mathbf{A}) \\ &= \log N! + N\Big[\phi_N(\boldsymbol{\beta}) - \sum_{\ell=3}^{K} \beta_\ell \frac{\partial\phi_N(\boldsymbol{\beta})}{\partial\beta_\ell}\Big] \\ &= N\log(N)\Big(1 - \sum_{\ell=3}^{K}\frac{m_\ell}{\ell}\Big) + \mathcal{O}(N). \end{aligned} \tag{2.23}$$

Since $\sum_{\ell=3}^{K}(m_\ell/\ell) \leq \frac{1}{3}\sum_{\ell=3}^{K} m_\ell \leq \frac{1}{3}$, the leading order will for large N always scale as $N\log(N)$, and be bounded according to $S_N \geq \frac{2}{3}N\log(N) + \mathcal{O}(N)$, but with a reduced prefactor if we increase the fraction of nodes in short loops. The lower bound is achieved in the disconnected phase, when $m_3 = 1$ and $m_{\ell>3} = 0$.

2.2.4 Spectral densities of adjacency matrices

A graph can be represented uniquely by its adjacency matrix $\{A_{ij}\}$, where $A_{ij} \in \{0,1\}$, and $A_{ij} = 1$ if and only if there is a link from j to i. The set $\mathcal{G}_N$ contains only simple nondirected graphs, so our adjacency matrices are symmetric and with zero diagonal elements. The eigenvalue density of the adjacency matrix of a graph $\mathbf{A}$,

$$\varrho(\mu|\mathbf{A}) \;=\; \frac{1}{N}\sum_{i=1}^{N}\delta[\mu-\mu_i(\mathbf{A})], \tag{2.24}$$

contains valuable information on the statistics of loops in the graph. Here the sum runs over the set of (real) eigenvalues $\{\mu_i(\mathbf{A})\}_{i=1,\ldots,N}$ of $\mathbf{A}$, taking into account multiplicities. For instance, the number of closed paths in $\mathbf{A}$ is proportional to $\int \mathrm{d}\mu\; \varrho(\mu|\mathbf{A})\mu^\ell$. Our main quantity of interest will be the expected density, averaged over the ensemble probabilities (2.1), in the infinite size limit,

$$\varrho(\mu) = \lim_{N\to\infty}\sum_{\mathbf{A}\in\mathcal{G}_N} p(\mathbf{A})\varrho(\mu|\mathbf{A}). \tag{2.25}$$

The adjacency matrix of a graph that consists of a single loop of length ℓ has the Toeplitz form, and is therefore diagonalized trivially, leading to the density

$$\varrho_\ell(\mu) \;=\; \frac{1}{\ell}\sum_{r=0}^{\ell-1}\delta\Big(\mu-2\cos(2\pi r/\ell)\Big). \tag{2.26}$$

If the loop length ℓ diverges, this density becomes continuous (in a distributional sense), see e.g. [46],

$$\varrho_\infty(\mu) \;=\; \lim_{\ell\to\infty}\varrho_\ell(\mu) = \frac{1}{\pi}\frac{\theta(2-|\mu|)}{\sqrt{4-\mu^2}}. \tag{2.27}$$

The set of eigenvalues for each $\mathbf{A}\in\mathcal{G}_N$ will just be the union of all the sets of eigenvalues of the disjoint loops of which it is composed, taking multiplicities into account:

$$\begin{aligned}\varrho(\mu|\mathbf{A}) \;&=\; \frac{1}{N}\sum_{\ell=3}^{N} n_\ell(\mathbf{A})\sum_{r=0}^{\ell-1}\delta\Big(\mu-2\cos\Big(\frac{2\pi r}{\ell}\Big)\Big)\\ &=\; \sum_{\ell=3}^{N}\frac{\ell n_\ell(\mathbf{A})}{N}\varrho_\ell(\mu).\end{aligned} \tag{2.28}$$

Upon averaging over the ensemble, using (2.3) and our earlier observation that for $N\to\infty$

the fraction of nodes in loops of finite length $L > K$ vanishes, we immediately obtain the asymptotic ensemble-averaged spectrum corresponding to (2.1), expressed in terms of (2.26) and (2.27):

$$\varrho(\mu) \;=\; \sum_{\ell=3}^{K} m_\ell \varrho_\ell(\mu) + m_\infty \varrho_\infty(\mu). \tag{2.29}$$

Since we are working with regular graphs, we can immediately recover the spectrum of the Laplacian operator ($\mathbf{L} = 2\mathbf{I} - \mathbf{A}$) by the change of variable $\mu \to 2 - \lambda$.

2.3 Grand Canonical approach

Within the canonical approach one finds that, if N is sufficiently large, graphs generated randomly from (2.1) will all display the same values of the main intensive quantities, such as the fraction of ℓ-loops (modulo finite size fluctuations). We expect a similar claim to hold if we sample randomly both the graphs and the number N of nodes, i.e. if we work with grand canonical graph ensembles. The grand partition function of our ensemble with weights $w_N = \mathrm{e}^{-\mu N}/N!$ (where $\mu > 0$) is given by

$$Q(\boldsymbol{\beta}) = \sum_{N=1}^{\infty} w_N Z_N(\boldsymbol{\beta}), \tag{2.30}$$

with $Z_N(\boldsymbol{\beta})$ defined in (2.2). The divisor $N!$ in w_N will simplify our calculation, without losing the benefits of the thermodynamic limit (since we will find that for $\mu \to 0$ the expected system size still diverges). Direct calculation of $Q(\boldsymbol{\beta})$ now circumvents the integration over ω:

$$\begin{aligned}
Q(\boldsymbol{\beta}) \;&=\; \sum_{\boldsymbol{n}} \Big(\prod_{\ell=3}^{\infty} \frac{\mathrm{e}^{-\mu \ell n_\ell}}{(2\ell)^{n_\ell} n_\ell!} \Big) \left(\prod_{\ell=3}^{K} \mathrm{e}^{\ell \beta_\ell n_\ell} \right) \\
&=\; \left[\prod_{\ell=3}^{K} \left(\sum_{n \geq 0} \frac{\mathrm{e}^{(\beta_\ell - \mu)\ell n}}{(2\ell)^n n!} \right) \right] \left[\prod_{\ell > K} \left(\sum_{n \geq 0} \frac{\mathrm{e}^{-\mu \ell n}}{(2\ell)^n n!} \right) \right] \\
&=\; \exp\left(\sum_{\ell=3}^{K} \frac{\mathrm{e}^{(\beta_\ell - \mu)\ell}}{2\ell} + \sum_{\ell > K} \frac{\mathrm{e}^{-\mu \ell}}{2\ell} \right) \\
&=\; \exp\left(\sum_{\ell=3}^{K} \frac{\mathrm{e}^{(\beta_\ell - \mu)\ell}}{2\ell} - \frac{1}{2}\log(1 - \mathrm{e}^{-\mu}) - \sum_{\ell=1}^{K} \frac{\mathrm{e}^{-\mu \ell}}{2\ell} \right),
\end{aligned} \tag{2.31}$$

where we used $\sum_{\ell>0} x^\ell/\ell = -\log(1–x)$. From $Q(\boldsymbol{\beta})$ we obtain, in turn, the grand potential $\Omega(\boldsymbol{\beta}) = -\log Q(\boldsymbol{\beta})$:

$$\Omega(\boldsymbol{\beta}) \;=\; \sum_{\ell=1}^{K} \frac{\mathrm{e}^{-\mu\ell}}{2\ell} - \sum_{\ell=3}^{K} \frac{\mathrm{e}^{(\beta_\ell-\mu)\ell}}{2\ell} + \frac{1}{2}\log(1-\mathrm{e}^{-\mu}). \tag{2.32}$$

Its partial derivatives with respect to μ and $\boldsymbol{\beta}$ yield the average system size, via $\langle N\rangle = \partial\Omega(\boldsymbol{\beta})/\partial\mu$, and the average number of length-ℓ loops (for $\ell = 3,\ldots,K$), via $\langle n_\ell(\mathbf{A})\rangle = -\ell^{-1}\partial\Omega(\boldsymbol{\beta})/\partial\beta_\ell$. We thereby find that

$$\begin{aligned}\langle N\rangle \;&=\; \frac{1}{2}\frac{\mathrm{e}^{-\mu}}{1-\mathrm{e}^{-\mu}} + \frac{1}{2}\sum_{\ell=3}^{K}\mathrm{e}^{(\beta_\ell-\mu)\ell} - \frac{1}{2}\sum_{\ell=1}^{K}\mathrm{e}^{-\mu\ell} \\ &=\; \frac{1}{2}\frac{\mathrm{e}^{-\mu(K+1)}}{1-\mathrm{e}^{-\mu}} + \frac{1}{2}\sum_{\ell=3}^{K}\mathrm{e}^{(\beta_\ell-\mu)\ell}\end{aligned} \tag{2.33}$$

and

$$\langle n_\ell(\mathbf{A})\rangle \;=\; \frac{1}{2\ell}\mathrm{e}^{(\beta_\ell-\mu)\ell}. \tag{2.34}$$

Clearly, $\langle N\rangle$ diverges for $\mu \to 0$, which gives our thermodynamic limit. In this limit we can then work out for $\ell \in \{3,\ldots,K\}$ the ratios

$$\lim_{\mu\to 0}\frac{\ell\langle n_\ell\rangle}{\langle N\rangle} \;=\; \lim_{\mu\to 0}\frac{\mathrm{e}^{(\beta_\ell-\mu)\ell}}{\frac{\mathrm{e}^{-\mu(K+1)}}{1-\mathrm{e}^{-\mu}} + \sum_{\ell'=3}^{K}\mathrm{e}^{(\beta_{\ell'}-\mu)\ell'}} = 0. \tag{2.35}$$

Similar to the canonical case, any μ-independent $\boldsymbol{\beta}$ will asymptotically always yield a vanishing fraction of nodes in loops of length $\ell \leq K$. It is clear from (2.34) that without a re-parametrization, the expected value of ℓ-loops only increases exponentially with β_ℓ in the thermodynamic limit. We need to re-parametrize in such a way that the expected number of ℓ-loops increases as the expected system size increases. The re-parametrization required to obtain a nontrivial thermodynamic limit is $\beta_\ell = \tilde{\beta}_\ell + \ell^{-1}\log\langle N\rangle$. Upon following this prescription, we then reproduce the canonical result

$$\frac{\ell\langle n_\ell\rangle}{\langle N\rangle} \;=\; \frac{1}{2}\mathrm{e}^{(\tilde{\beta}_\ell-\mu)\ell}, \tag{2.36}$$

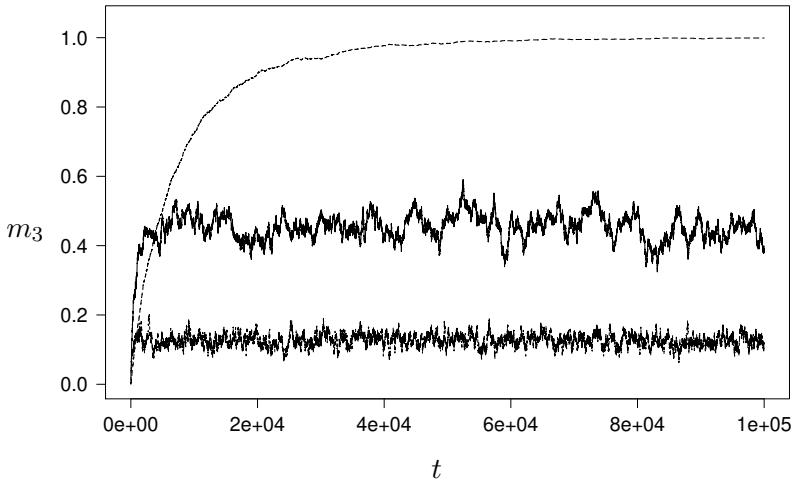


Figure 2.1: Examples of the evolution of the fraction m_3 of nodes in triangles, measured during MCMC simulations, for $K = 3$. Time is defined as the number of accepted edge swap moves per link. The bottom two curves correspond to the connected phase of the ensemble, equilibrating to the values $m_3 = 0.125$ for $\tilde{\beta}_3 = \frac{1}{3}\log(0.25)$, and to $m_3 = 0.45$ for $\tilde{\beta}_3 = \frac{1}{3}\log(0.9)$. The top curve corresponds to the disconnected phase, here the MCMC process is equilibrating to the value $m_3 = 1$.

and our expression (2.33) for $\langle N\rangle$ now becomes

$$\langle N\rangle \;=\; \frac{\mathrm{e}^{-\mu(K+1)}}{(1-\mathrm{e}^{-\mu})\big(2-\sum_{\ell=3}^{K}\mathrm{e}^{(\tilde{\beta}_\ell-\mu)\ell}\big)}. \tag{2.37}$$

The re-parametrization of $\boldsymbol{\beta}$ now depends on $\boldsymbol{\beta}$ itself, via $\langle N\rangle$, and has to be consistent with a nonnegative value for (2.37), i.e. with $\frac{1}{2}\sum_{\ell=3}^{K}\mathrm{e}^{(\tilde{\beta}_\ell-\mu)\ell} \leq 1$. Expression (2.36) gives us the physical interpretation $\sum_{\ell=3}^{K}\ell\langle n_\ell\rangle/\langle N\rangle \leq 1$. In the limit $\mu \to 0$ the condition becomes $\sum_{\ell=3}^{K}\mathrm{e}^{\ell\tilde{\beta}_\ell} \leq 2$. In the case of inequality we have $\sum_{\ell=3}^{K}\ell\langle n_\ell\rangle/\langle N\rangle < 1$, so we are in the connected phase. The case of equality reproduces our earlier phase transition condition (2.21) and we enter the disconnected phase; here the thermodynamic limit is reached already for nonzero μ, and we can again recover our canonical equations, with $\exp(-\mu)$ now playing the role of the canonical order parameter x.

2.4 Numerical simulations

Calculating $Z_N(\boldsymbol{\beta})$ by numerical enumeration for large values of N is not a realistic option, since the size of the set $\mathcal{G}_N$ grows super-exponentially with N. Instead, to test our theoretical predictions we have sampled graphs from the ensemble (2.1) using the Markov Chain Monte Carlo (MCMC) method described in appendix A, presented in [6, 9]. Starting

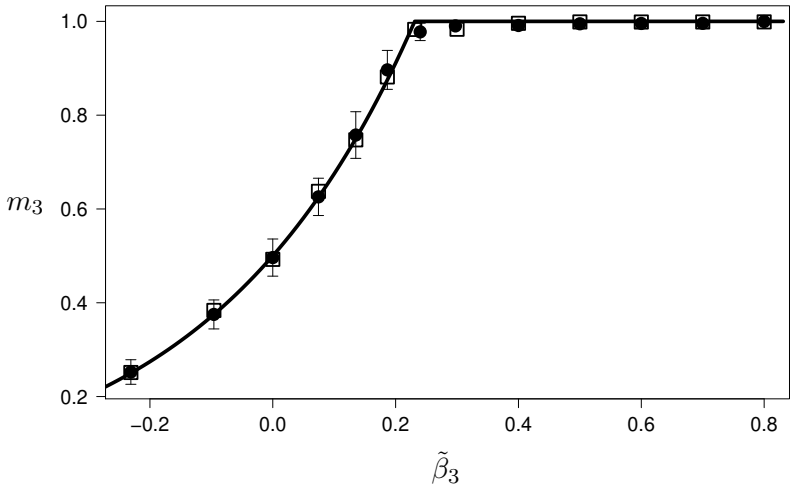


Figure 2.2: Values of m_3 shown versus $\tilde{\beta}_3$ for ensembles with $K = 3$. Numerical results, measured upon equilibration of the MCMC processes, are shown as black dots with error bars for $N = 1000$, and as squares for $N = 5000$ (error bars for $N = 5000$ are not shown; their sizes are similar to or smaller than the squares). The solid line is the prediction of (2.38).

from an arbitrary 2-regular N-node graph, this stochastic process is based on executing repeated (degree-preserving) edge swap moves with appropriate nontrivial move acceptance probabilities, constructed such that the Markov chain's equilibrium distribution is the target measure (2.1). In each simulation experiment, the MCMC process was first run for 10^5 to 10^6 accepted moves per link. After this randomization stage, the instantaneous state $\mathbf{A}$ arrived at by the chain was defined to be our graph sample. We have limited our simulations to ensembles with $K = 3$ and $K = 4$. The degree of equilibration achieved by the MCMC during a run of 10^5 accepted moves per link is illustrated in Figure 2.1, where we show typical evolution curves of the order parameter m_3 during the stochastic process.

For $K = 3$ we have just one control parameter $\tilde{\beta}_3$, and the order parameter is the fraction m_3 of nodes in triangles. The theory claims that, for large N, the graphs from our ensemble will be collections of triangles and large rings. The key equations (2.20,2.21,2.22) reduce to the following predictions, with $\tilde{\beta}_c = \frac{1}{3}\log(2) \approx 0.23105\ldots$:

$$\begin{aligned} &\tilde{\beta}_3 < \tilde{\beta}_c: \quad m_3 = \tfrac{1}{2}\mathrm{e}^{3\tilde{\beta}_3}, \quad \textit{connected phase}, \\ &\tilde{\beta}_3 > \tilde{\beta}_c: \quad m_3 = 1, \qquad \textit{disconnected phase}. \end{aligned} \tag{2.38}$$

Numerical simulations with sizes $N = 1000$ and $N = 5000$ show excellent agreement with these predictions, as shown in Figure 2.2, both in terms of the values of m_3 and in terms

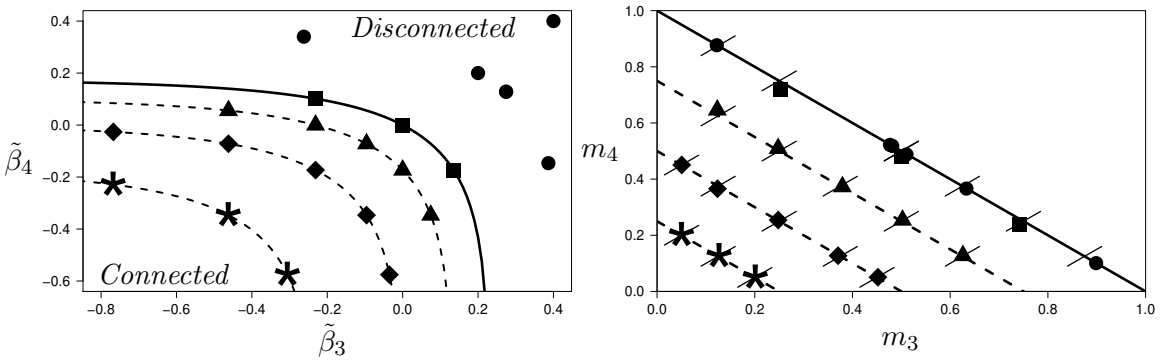


Figure 2.3: Left panel: the plane of control parameters for $K = 4$. The solid black line is the critical line $e^{3\tilde{\beta}_3} + e^{4\tilde{\beta}_4} = 2$ (here $m_\infty = 0$). The dashed lines correspond to parameter combinations with constant m_∞, taking the values $m_\infty \in \{0.75, 0.5, 0.25\}$, from bottom to top. The markers represent parameter combinations chosen for MCMC simulations. Right panel: the fractions (m_3, m_4) associated with the control parameter combinations in the left panel. Here the markers represent the simulation results, measured after execution of 10^4 accepted moves per node in the MCMC to secure equilibration. The results are indeed found on the respective lines predicted by the theory. Note that the theory predicts that all parameter combinations in the disconnected phase $e^{3\tilde{\beta}_3} + e^{4\tilde{\beta}_4} \geq 2$, should be mapped to the line $m_1 + m_2 = 1$ in the right panel. Error bars were omitted, as they are as big as or smaller than the markers.

of the location of the transition.

For $K = 4$ we have two control parameters, $\tilde{\beta}_3$ and $\tilde{\beta}_4$, and the theory claims that for large N the graphs from our ensemble will now be collections of triangles, squares and large rings. Here the key equations (2.20,2.21,2.22) predict that the transition line in parameter space is given by $e^{3\tilde{\beta}_3} + e^{4\tilde{\beta}_4} = 2$, and that the fractions m_3 and m_4 of nodes found in triangles and squares, respectively, are solved (together with the auxiliary order parameter x, in the disconnected phase) from:

$$\begin{aligned} e^{3\tilde{\beta}_3} + e^{4\tilde{\beta}_4} < 2: \quad & m_3 + m_4 < 1, \quad \textit{connected phase}, \\ & m_3 = \tfrac{1}{2} e^{3\tilde{\beta}_3}, \quad m_4 = \tfrac{1}{2} e^{4\tilde{\beta}_4}, \\ e^{3\tilde{\beta}_3} + e^{4\tilde{\beta}_4} > 2: \quad & m_3 + m_4 = 1, \quad \textit{disconnected phase}, \\ & m_3 = \tfrac{1}{2} x^3 e^{3\tilde{\beta}_3}, \quad m_4 = \tfrac{1}{2} x^4 e^{4\tilde{\beta}_4}. \end{aligned} \tag{2.39}$$

Figure 2.3 (left panel) shows the resulting predicted phase diagram in the $(\tilde{\beta}_3, \tilde{\beta}_4)$ plane. The mapping $(\tilde{\beta}_3, \tilde{\beta}_4) \mapsto (m_3, m_4)$ will map the lower region of the phase diagram (the connected phase) to the interior of the triangle $m_3 + m_4 < 1$ in the right panel of Figure 2.3. The upper region of the phase diagram on the left (the disconnected phase), including the

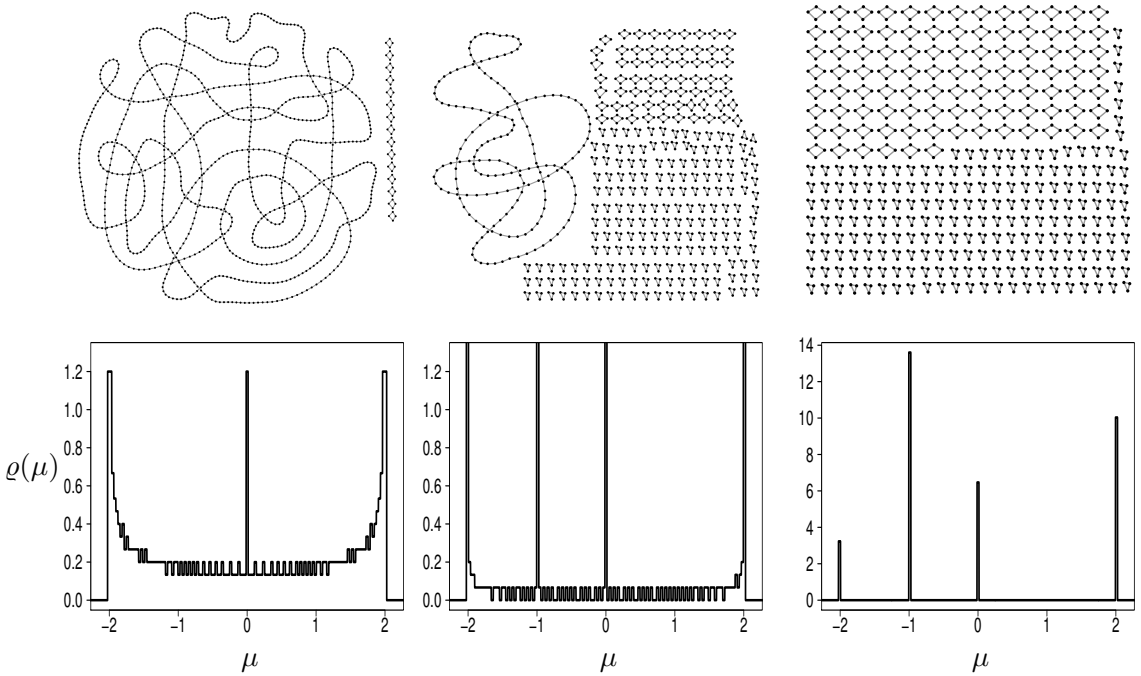


Figure 2.4: Top row: typical graphs sampled numerically via MCMC from the canonical ensemble (2.1) of 2-regular nondirected simple graphs, for $N = 1000$. Left: $(m_3, m_4) = (0.0, 0.06)$ and $m_\infty = 0.94$. Middle: $(m_3, m_4) = (0.25, 0.56)$ and $m_\infty = 0.19$. Right: $(m_3, m_4) = (0.39, 0.61)$ and $m_\infty = 0$. The bottom row shows the eigenvalue spectra of the corresponding three adjacency matrices, computed by direct numerical diagonalization. The locations of the peaks are seen to agree with the theoretical predictions of (2.29). Note the different scale in the third spectrum graph, to emphasize the weights of the δ-peaks.

critical line, will be mapped to the line $m_3 + m_4 = 1$ in the right panel. To test also these predictions against numerical simulations, we have chosen multiple points $(\tilde{\beta}_3, \tilde{\beta}_4)$ in both regions of the phase diagram, grouped such that the predicted values of $m_\infty = 1 - m_3 - m_4$ were always in the set $\{0.25, 0.5, 0.75\}$. The prediction would therefore be that in the (m_3, m_4) plane these groups of points should be found on the lines $m_3 + m_4 = 1 - m_\infty$. Upon measuring the fractions m_3 and m_4 via MCMC in the corresponding graph ensembles, these predictions are once more validated convincingly. See Figure 2.3.

Graphs sampled from our ensemble with $K = 4$ do indeed typically consist of controlled numbers of triangles and squares, and a long ring. Figure 2.4 shows examples of such graphs, obtained via MCMC, together with the eigenvalue spectra of their adjacency matrices (obtained by numerical diagonalization). Also the observed spectra agree with the corresponding theoretical predictions (2.29).

2.5 Discussion

In this chapter we presented an analytical solution for an exponential random graph ensemble with a controllable density of short loops. Whereas one would normally not expect such non-treelike graph ensembles to be solvable, here this is possible as a consequence of imposing a local degree constraint of strict 2-regularity. We found a second order phase transition, which separates a connected phase with large and small loops from a disconnected phase where the graphs are typically formed only of extensively many short loops. The short loops appear in controlled proportions, for which we found analytical expressions in terms of the ensemble's parameters. We also derived an analytical expression for the critical submanifold in the phase diagram, and for the expected eigenvalue spectrum of the graphs' adjacency matrices.

We analyzed both the canonical and the grand canonical formulation of the ensemble. The canonical version was solved via steepest descent integration. In the grand canonical version one avoids steepest descent integration, but (as always) the chemical potential takes over the role of the steepest descent integration variable of the canonical version. In the thermodynamic limit, the canonical and grand canonical routes result in identical equations. These equations are found to give highly accurate predictions already for modest graph sizes, such as $N = 1000$, as we confirmed in numerical simulations.

The parameter K represents the largest loop length that is controlled in our model. For $K = 3$ one controls only the number of triangles, and our ensemble becomes similar to that of Strauss [10–12] with average degree two. The remaining difference is that in the Strauss model the average degree is imposed implicitly via an overall 'soft' constraint, while in the present model all degree values are imposed as local 'hard' constraints. Due to this difference, the degeneration of the Strauss model to a phase where the complete clique has probability one (so the number of triangles can no longer be tuned) is avoided in the present ensemble. The complete clique is simply no longer an allowed configuration, and hence the number of triangles becomes fully tuneable, if the model parameters scale appropriately with the system size. In addition, in [14] it is shown that the 'soft' version of our model would have a phase diagram reminiscent of ours. In both cases the sign of a linear combination of functions of the parameters determines the phase of the ensemble. However, in the 'soft' case of [14] there is a transition from an almost ER-like phase to a clique, while our model exhibits tuneability of the densities in both phases.

The generalization of the present model to other degree distributions has been studied numerically for other degree distributions by other authors in the last decade, [24, 26]. Nevertheless in both studies there is no mention of any use of the correct nontrivial acceptance probabilities (A.5). This would in principle make their works incorrect, but as we will show in the next chapter the overall phenomenology observed agrees with what we found using the correct MCMC algorithm.

In our view, the main merit of the present model is that its analytical solution helps us understand more complicated 'loopy' graph ensembles. We are aware that the analytical route taken in this case is surely impossible for other models. Nevertheless, it provides an explicit analytical solution that reproduces the main features of non-treelike random graph ensembles with hard degree constraints. It helps us understand phenomenology that had so far only been studied numerically. The generalizations presented in the next chapters bear important resemblance to equations presented here, such as (2.10) where the loop series can be appreciated. It will also serve as a benchmark model against which more general solution strategies for non-treelike random graphs can be tested, such as [47], which deals with spectrally constrained maximum entropy graph ensembles. This is developed in Chapters 5 and 6. The moments of a graph's spectral density are related to its numbers of loops, via the traces of powers of the adjacency matrix. In fact, the present model is a special case of the family of ensembles studied in further chapters, from which it can be obtained by choosing 2-regular degrees and an appropriate polynomial functional Lagrange parameter. The analytical and numerical results of this chapter suggest that, to obtain phase transitions, the functional Lagrange parameters in spectrally constrained maximum entropy graph ensembles may need to have a specific scaling with the system size.

In the next chapters we will present the generalization of this ensemble to an arbitrary degree distribution. The phenomenology of the connected-disconnected transition will become richer but not drastically different. For the rest of this thesis we will work with biases of traces $\alpha_\ell \mathrm{Tr}(\mathbf{A}^\ell)$ instead of biases of loop counters $\beta_\ell \ell n_\ell(\mathbf{A})$. In general changing from β_ℓ to α_ℓ is a complicated task, except for the particular case of $\ell = 3$ where the mapping is simple, $\beta_3 = 2\alpha_3$.

Chapter 3

Random loopy graphs with an arbitrary degree distribution

In this chapter we introduce a general model of loopy random graphs. In this case we consider an ensemble of random graphs with an arbitrary degree biased with respect to the number of triangles.

3.1 The model

We will study a random graph ensemble defined on the set of N node graphs with a given degree distribution. Formally a graph is an ordered pair (V, E) of nodes and edges respectively. We will look at graphs through their adjacency matrices $\mathbf{A}$ defined as $(\mathbf{A})_{ij} = A_{ij} = 1$ if $(i, j) \in E$ and 0 otherwise. We will only be concerned with simple undirected graphs, which in terms of the adjacency matrix means they will all be symmetric and have zeros for all diagonal elements. The degree of a node is the number edges around it, $k_i(\mathbf{A}) = \sum_j A_{ij}$. Throughout this chapter we will work with graphs that have exactly the same degree sequence $\{k_i\}_{i=1,\ldots,N}$. Such sequences will be taken to be a sample of a given distribution $p(k)$. In the large N limit we know that the empirical distribution of degrees will closely match the target distribution, $p(k) = \lim_{N\to\infty} \sum_{i=1}^{N} N^{-1}\delta_{k,k_i}$.

In the matrix language, the number of triangles inside a graph is very simple to calculate provided we identify them with the loops of length three up to overcounting. We say that there is a loop of length three (ℓ_3) around node i if there exists j and k such that $(i, j) \in E$, $(j, k) \in E$, and $(k, i) \in E$. Since our graph is simple it follows that necessarily i, j, k are all different. In the language of the adjacency matrix the indicator function for

a given loop takes the following very simple form,

$$\mathbb{I}\left[(i \to j \to k \to i) \in \mathbf{A}\right] = A_{ij}A_{jk}A_{ki}. \tag{3.1}$$

The total number of 3-loops is then simply given by the trace of the third power of the adjacency matrix,

$$\mathcal{M}(\mathbf{A}) = \sum_{i,j,k} \mathbb{I}\left[(i \to j \to k \to i) \in \mathbf{A}\right] = \mathrm{Tr}(\mathbf{A}^3). \tag{3.2}$$

Following the discussion from the introduction, we now define an ensemble of random graphs such that the average number of triangles can be controlled using a parametrized distribution over graphs denoted by $p(\mathbf{A})$. Our choice is a maximum entropy (ME) ensemble. That is, we take $p(\mathbf{A})$ to be such that the average number of triangles is fixed,

$$\mathcal{M}^* = \sum_{\mathbf{A}} p(\mathbf{A})\mathrm{Tr}(\mathbf{A}^3), \tag{3.3}$$

and such that the degree sequence $\{k_i\}_{i=1,\dots,N}$ is followed exactly. Among all the possible distributions $p(\mathbf{A})$ that share these two properties we choose the one that maximizes the Shannon entropy, $S[p] = -\sum_{\mathbf{A}} p(\mathbf{A}) \log p(\mathbf{A})$. This will guarantee that the distribution is statistically unbiased, [8].

The ME distribution is an exponential with one tuneable parameter α,

$$p(\mathbf{A}) = \frac{1}{Z(\alpha)} \mathrm{e}^{\alpha \mathrm{Tr}(\mathbf{A}^3)} \prod_{i=1}^{N} \delta_{k_i, \sum_j A_{ij}}. \tag{3.4}$$

The product over deltas enforces the degree sequence of the graph. Notice that for $\alpha = 0$ the ensemble is simply the configuration model (CM), a uniform distribution over all graphs with degree sequence $\mathbf{k} = (k_i)_{i=1,\dots,N}$.

Our main observable of the ensemble will be the number of 3-loops *per node*, making it comparable among different sizes. We will refer to it as the *loop density*,

$$m(\alpha) = \left\langle \frac{1}{N}\mathcal{M}(\mathbf{A}) \right\rangle = \left\langle \frac{1}{N}\mathrm{Tr}(\mathbf{A}^3) \right\rangle. \tag{3.5}$$

Where $\langle f(\mathbf{A}) \rangle = \sum_{\mathbf{A}} p(\mathbf{A}) f(\mathbf{A})$. This quantity reflects the typical number of loops in the neighborhood of node. Each node can have a different maximum number of triangles

depending on its degree k_i. Once a random graph ensemble like (3.4) is defined it is desirable to have both an algorithm to generate samples numerically and an analytic theory of its statistical properties.

In order to generate samples from an exponential ensemble such as (3.4) we chose to use a Markov Chain Monte Carlo (MCMC) approach. The algorithm consists of starting with a seed graph satisfying the degree sequence and evolving it by performing degree preserving edge swaps as the one shown in Figure A.1. Not all edge swaps are performed, they are either accepted or rejected with a given nontrivial acceptance probability that not only takes into account the specific ensemble (3.4) but that also takes into account the availability of possible edge swaps as the graph evolves. The theory of this MCMC algorithm was developed in [9] and pedgagogically presented in the book [6]. We include a summary in Appendix A.

To find an analytic expression for the loop density we need to calculate the following generating function.

$$\phi(\alpha) = \frac{1}{N}\log Z(\alpha) = \frac{1}{N}\log \sum_{\mathbf{A}} \mathrm{e}^{\alpha \mathrm{Tr}(\mathbf{A}^3)} \prod_{i=1}^{N} \delta_{k_i, \sum_j A_{ij}}, \tag{3.6}$$

$$m(\alpha) = \frac{\partial \phi}{\partial \alpha} = \left\langle \frac{1}{N}\mathrm{Tr}(\mathbf{A}^3) \right\rangle. \tag{3.7}$$

Ideally the knowledge of functions $\phi(\alpha)$ and $m(\alpha)$ would allow us to generate random graphs with any desired loop density. Although it is not possible to calculate (3.6) analytically, in section 3.3 we will show that a small α approximation will give very good results for a wide range of values. Additionally we will give a description of the general behavior of this ensemble for the whole range of α.

It turns out that another very important observable of the ensemble is the amount of *interaction* between the triangles in the graph. By interaction we mean the number of edges and nodes that different triangles share. The level of interaction varies in a nontrivial unexpected way for different values of α and different system sizes N. In order to measure the degree of interaction between loops, we define the following quantity,

$$r(\mathbf{A}) = \frac{\#\text{nodes in triangles}}{\#\text{ of triangles}} = 6\frac{\sum_{i=1}^{N} \Theta[(\mathbf{A}^3)_{ii}]}{\mathrm{Tr}(\mathbf{A}^3)} \in [0, 3]. \tag{3.8}$$

Where $\Theta(x) = 1$ if $x > 0$ and zero otherwise. This is the ratio of triangle vertices to triangles, it is independent of the total number of triangles in the graph. If $r(\mathbf{A}) = 3$ it

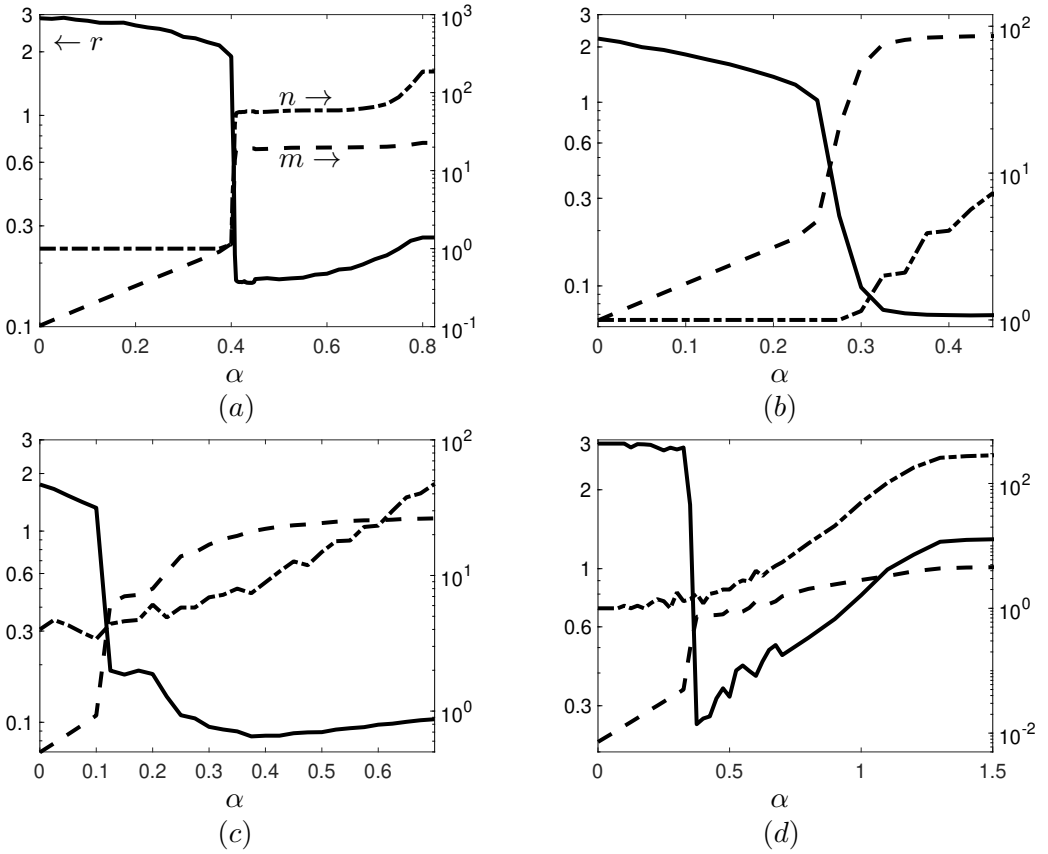


Figure 3.1: Plots for: $r(\alpha)$, loop interaction (3.8) - solid line, values on left axis; $m(\alpha)$, loop density - dashed line, values on right axis; $n(\alpha)$, number of connected components - dashed dotted line, right axis. (a) $p(k) = bim(k|3,7)$, (b) $p(k) = Poiss(k|10)$, (c) $p(k) = exp(k|4)$, (d) $p(k) = PL(k)$. Size $N = 1000$ in all cases. Error bars are omitted for clarity, see figures 3.5, 3.6 for examples.

means that regardless of the number of triangles, they are all noninteracting in the sense that they do not share edges or nodes. As $r(\mathbf{A})$ goes below 3 it means that they are sharing either edges or nodes. Some simple examples are shown on the top row of Figure 3.5. In the particular case when graphs form cliques of $q+1$ nodes we have $r(\mathbf{A}) = 6/(q^2-q)$, this is a natural lower bound for graphs of maximum degree q.

3.2 Main results

We are now going to outline the main results of our analysis of the ensemble (3.4).

We found the same initial behavior for all degree distributions for initial values of α, that is as it is increased from $\alpha = 0$. The main quantities defining the initial behaviour of the ensemble are the first two moments of the degree distribution, $c = \overline{k}$ and $\overline{k^2}$, where $\overline{f(k)} = \sum_{i=1}^{N} N^{-1} f(k_i)$. For the case of bounded degree distributions the maximum degree,

$q = \max_{i=1,\dots,N}\{k_i\}$, will also be of importance. Even though we consider only the case of finite system size, N, we will assume N is sufficiently large, at least $Np(q) >> q+1$, so that all degrees are represented with an extensive number of nodes.

In Figure 3.1 we can see that the loop density is as expected a growing function of α. We can also see that there are clearly different regimes, as had already been observed for regular graphs in [48]. Interestingly, to really understand the nature of the different regimes of the ensemble it is necessary to also look at two other graph observables: the level of interaction between loops, measured with $r(\mathbf{A})$ as defined in (3.8), and the number of connected components of the graph, $n(\mathbf{A})$. We define their respective ensemble averages, $r(\alpha) = \langle r(\mathbf{A})\rangle$ and $n(\alpha) = \langle n(\mathbf{A})\rangle$.

The observed regimes are the following ones.

- $\alpha \in [0, \alpha_1(N)]$: connected regime

  The behaviour of $m(\alpha)$ is exponential, following

  $$m(\alpha) = \frac{1}{N}\left(\frac{\overline{k^2}}{c} - 1\right)^3 \mathrm{e}^{6\alpha}. \tag{3.9}$$

  The only quantities that depend on $p(k)$ are the proportionality constant and the transition point $\alpha_1(N)$. This explicit and general formula is one of the main results of this work. It allows for an explicit calculation of α given a desired loop density, simply by inverting (3.9). For $\alpha = 0$ it is the expected loop density for large graphs, [5]. The degree of interaction between loops is as low as $r(\mathbf{A}) \sim 3$ for large graphs. The number of components of the graph is the same as in the $\alpha = 0$ case. It is relatively easy to obtain samples in this regime with the MCMC edge swap dynamics.

- $\alpha \in [\alpha_1(N), \alpha_2(N)]$: clustered regime

  We see that the triangle density $m(\alpha)$ grows faster than what is predicted by (3.9), making this formula inaccurate in this regime. Depending on the degree distribution the growth can have sudden jumps or it can be more smooth. The main difference with the previous regime is that loops start sharing edges. This can be seen since $\langle r(\mathbf{A})\rangle$ drops quickly from values around 3 to lower values. Nodes start to form clusters of similar degree. We call this the clustered phase and $\alpha_1(N)$ the clustering transition point.

- $\alpha \in [\alpha_2(N), \infty)$: disconnected regime

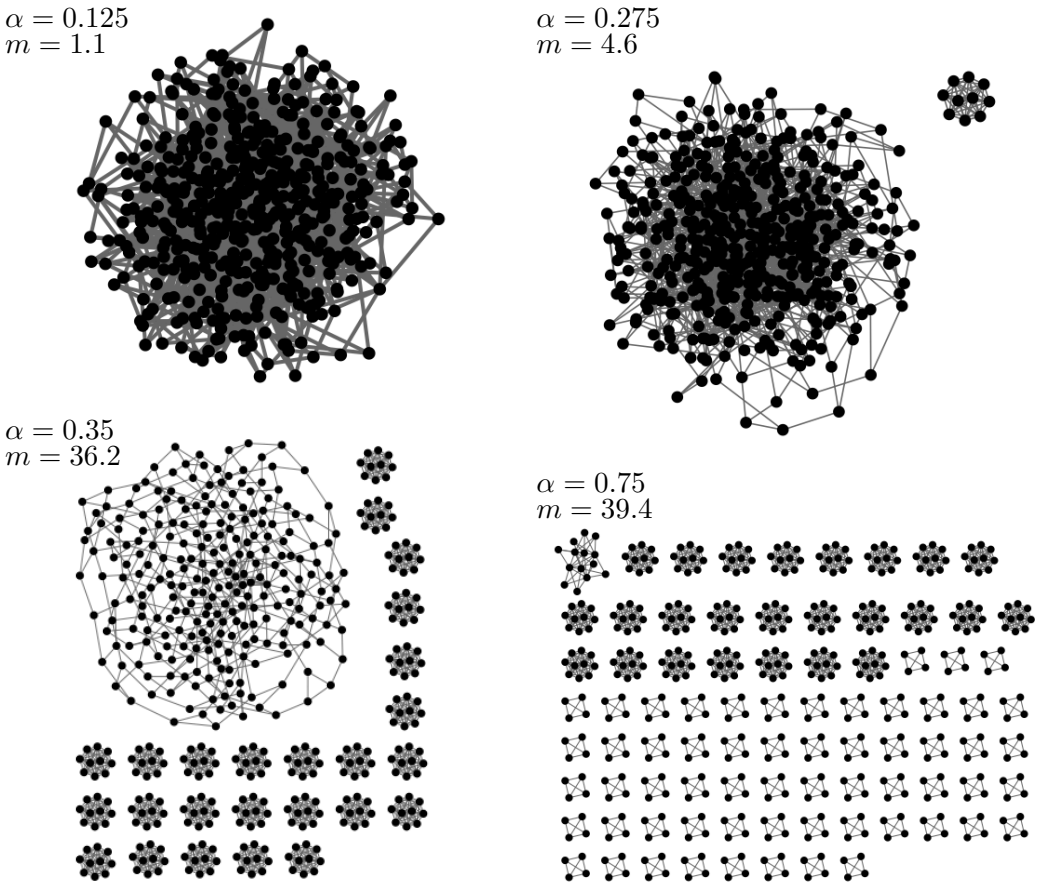


Figure 3.2: Results of numerical sampling simulations of the random graph ensemble (3.4), with $N = 500$ and $p(k) = \frac{1}{2}\delta_{k,3} + \frac{1}{2}\delta_{k,9}$. The four different images correspond to four different values of α, with different loop densities $m(\alpha)$, as indicated.

There is a topological change associated with the transition at $\alpha_2(N)$. The graph breaks down beyond the typical number of connected components of the corresponding CM. Small disconnected cliques of $k+1$ nodes will maximize the number of loops around a node of degree k, see Figures 3.2 and 3.3. Cliques of the maximum degree q will appear first and next the lower degrees. Even if due to finite size effects there are not enough nodes to generate cliques the graphs break down into small incomplete cliques. For this reason we call $\alpha_2(N)$ the *shattering* transition, and $[\alpha_2(N), \infty)$ the disconnected phase or shattered phase. The rest of the nodes will continue to be connected and follow qualitatively similar regimes but now for a new degree distribution that excludes the separated nodes.

We make the distinction between bounded and unbounded degree distributions $p(k)$ because it affects the way in which the ensemble behaves with N, for example in the

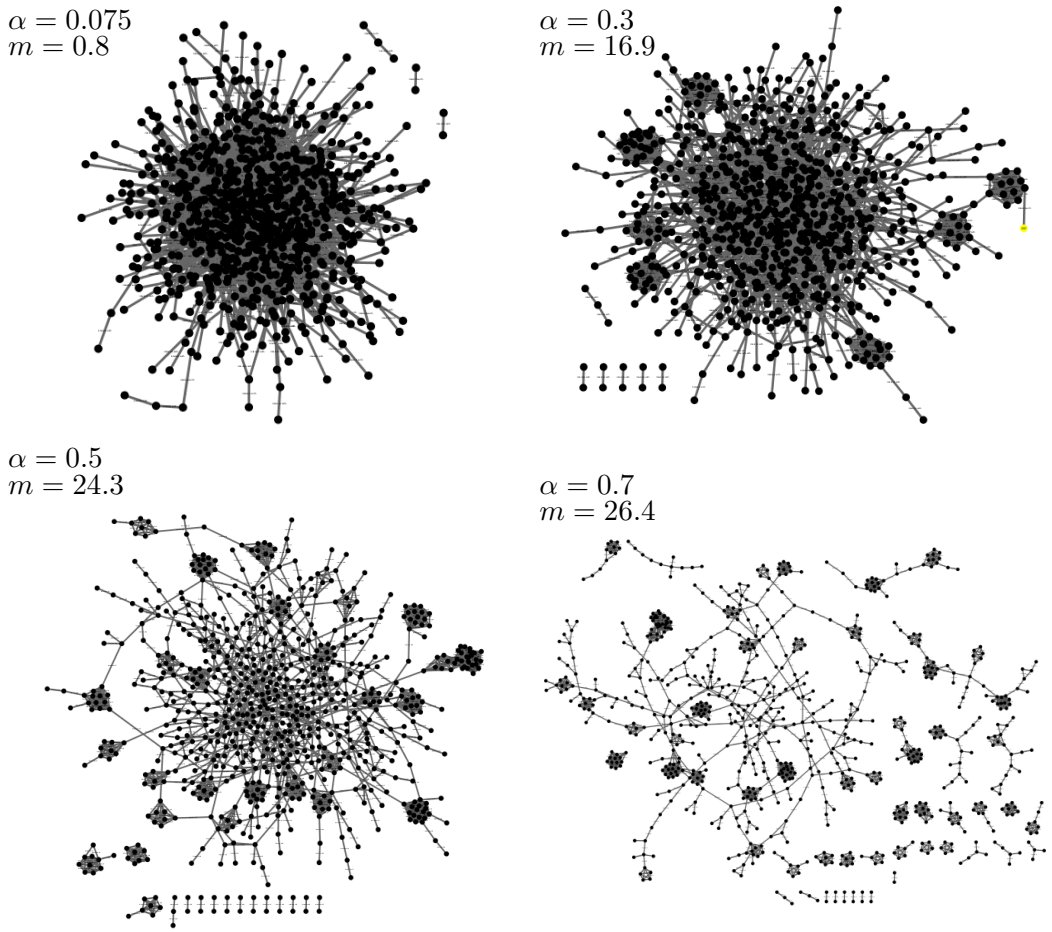


Figure 3.3: Results of numerical sampling simulations of the random graph ensemble (3.4), with $N = 1000$ and $p(k) = \frac{1}{4}\left(\frac{4}{5}\right)^k$. The four different images correspond to four different values of α, with different loop densities $m(\alpha)$, as indicated. Zero degree nodes are omitted.

asymptotics of $\alpha_1(N)$ and $\alpha_2(N)$. For the case of large graphs with bounded degree distributions the behavior is quite clear. In this case both transitions are quite close, $\alpha_1(N) \approx \alpha_2(N)$. There is a sudden appearance of disconnected cliques of $q+1$ nodes giving rise to a very sharp jump in $m(\alpha)$. There is strong numerical evidence and mathematical arguments to argue that both $\alpha_1(N)$ and $\alpha_2(N)$ scale as $\mathcal{O}(\log N)$ for large N. In the case of graphs with unbounded $p(k)$ the maximum degree will diverge slowly with N, therefore there are not many nodes of large degree to create cliques so the structures created when the graphs shatter are not so clear. We believe that the asymptotics of $\alpha_1(N)$ and $\alpha_2(N)$ should depend heavily on the tail of $p(k)$ as this is what governs the growth of the maximum degree with N. Nevertheless, in both cases the ensemble will end in a set of disconnected cliques as this is the graph that maximizes the number of loops around each node. The difference can be clearly seen when comparing Figure 3.3 and Figure 3.2, were different

degree distributions are chosen. For the graph with a bimodal degree distribution in Figure 3.3 the cliques appear immediately as the graph clusters, while for the one with an exponential distribution in Figure 3.2 one can see clusters appearing before the breaking down of the graph.

We point out that we will avoid speaking of transitions in the typical sense of statistical physics, since we are assuming very large N but not taking N to ∞. For this reason, the definitions given for $\alpha_1(N)$ and $\alpha_2(N)$ are rather of a descriptive nature, marking the values when $\langle r(\mathbf{A})\rangle$ drops for the first time and when $\langle n(\mathbf{A})\rangle$ increases for the first time, respectively. It will become apparent in the following section, the systems size N affects severely the ensemble.

In general, dealing analytically with expressions like (3.6) is very hard, especially for an arbitrary value of N. A typical approach is to derive exact results in the $N \to \infty$ and then show these are also good approximations for finite N. In contrast, in our case we found that it was necessary *not* to take the limit $N \to \infty$ strictly, but rather to work with asymptotically vanishing expressions for the loop density, $m(\alpha) = \mathcal{O}\left(N^{-\delta}\right)$. The most important example is that of the *connected noninteracting loopy regime*. Here equation (3.9) shows that $m(\alpha) = 0$ asymptotically, while just reporting this would be correct, it is the way in which it approaches 0 that gives us formula (3.9) which is seen to be very accurate. One would typically scale α with N to avoid this effect, but it will become clear that in that case $m(\alpha_1(N)) \to 0$ for any proper scaling of α with N, meaning the description of the first regime would vanish, which is not something that we want.

Regarding the MCMC sampling we point out that convergence from a given random initial condition into equilibration takes increasingly more edge swaps as α is increased, which is to be expected. Only for values of α in the connected regime, $\alpha \in [0, \alpha_1(N))$, does equilibration occur quickly enough to sample in a reasonable amount of time on a personal computer. As one might expect, close to the transitions there is a very significant divergence of relaxation times. We conjecture that the main reason for this dynamical change is precisely the clustering of triangles. In order to break a clique one has to destroy many triangles, an event that becomes extremely unlikely to be observed in the dynamics for large graphs. Therefore we expect there to be an effective ergodicity breaking when sampling with MCMC for $\alpha > \alpha_2(N)$ and large N.

For this reason, from the point of view of applied network science, working with our loopy ensemble (3.4) is easy but it has to be done carefully. We can summarize all previous

points by saying, that given a network, it is possible to randomize it through edge swaps while keeping the same loop density. But there will be two different problems. First, it could be that it takes a long time to sample correctly, and second it could be that the samples generated have a completely different topology, while still sharing the same loop density as seen through the value of $r(\mathbf{A})$. The first problem is a matter of computing power and speed, which can be eventually alleviated because the power of computers is constantly improving. The second problem is more difficult and essentially unsolvable without modifying (3.4). It is just a hard fact that if the graph you want to randomize has a different value of $r(\mathbf{A})$ to the corresponding value of $\langle r(\mathbf{A})\rangle$ with the same loop density, then all the samples will be typically very different in structure even though they share the same loop density.

In the following sections we will give analytic approximations and estimates for the density in the first regime, and an estimate on when the relevant transitions occur. It is important to note that both transitions have a clear system size dependence. The scaling of the different quantities with N will also be presented when possible.

3.3 The connected regime

We will now present a tractable approximation for the generating function (3.6). It is analogous to the one presented in [48], but generalized for an arbitrary degree distribution $p(k)$ with finite first and second moments.

We use a small α (or large N) approximation in order to derive (3.9) using a known result about the distribution of triangles in the CM. Amazingly it gives very good results, suggesting it could be exact asymptotically, at least for bounded degree distributions. We note that the number of 3-loops corresponds to six times the number of triangles. If we denote by $T(\mathbf{A})$ the number of triangles in $\mathbf{A}$ we get

$$\mathrm{Tr}(\mathbf{A}^3) = 6T(\mathbf{A}). \tag{3.10}$$

We can then calculate the generating function (3.6) in the following way.

$$\phi(\alpha) = \frac{1}{N}\log\sum_{\mathbf{A}} \mathrm{e}^{6\alpha T(\mathbf{A})}\prod_{i=1}^{N}\delta_{k_i,\sum_j A_{ij}} = \frac{1}{N}\log\sum_{T} \mathrm{e}^{6\alpha T} P_N(T) + \frac{1}{N}\log\mathcal{N}_{\mathbf{k}}. \tag{3.11}$$

Where we have introduced,

$$P_N(T) = \frac{1}{\mathcal{N}_{\mathbf{k}}} \sum_{\mathbf{A}} \delta_{T,T(\mathbf{A})} \prod_{i=1}^{N} \delta_{k_i,\sum_j A_{ij}}, \tag{3.12}$$

$$\mathcal{N}_{\mathbf{k}} = \sum_{\mathbf{A}} \prod_{i=1}^{N} \delta_{k_i,\sum_j A_{ij}}. \tag{3.13}$$

The main approximation now consists of replacing $P_N(T)$ by the known asymptotic distribution of isolated triangles, that is triangles that do not share edges or nodes. This has been rigorously established to be Poissonian in [5].

$$P_N(T) \sim Poiss(T|\lambda_t) = \mathrm{e}^{-\lambda_t} \frac{(\lambda_t)^T}{T!}, \tag{3.14}$$

$$\lambda_t = \frac{1}{6} \left(\frac{\sum_{i=1}^{N} k_i(k_i-1)}{\sum_{i=1}^{N} k_i} \right)^3 = \frac{1}{6} \left(\frac{\overline{k^2}}{c} - 1 \right)^3. \tag{3.15}$$

Using the analytical formula for the Poissonian distribution we can approximate (3.11),

$$\phi(\alpha) \approx \frac{1}{N} \lambda_t \left(\mathrm{e}^{6\alpha} - 1 \right) + \frac{1}{N} \log \mathcal{N}_{\mathbf{k}}, \tag{3.16}$$

$$m(\alpha) \approx \frac{1}{N} 6 \lambda_t \mathrm{e}^{6\alpha} = \frac{1}{N} \left(\frac{\overline{k^2}}{c} - 1 \right)^3 \mathrm{e}^{6\alpha}. \tag{3.17}$$

This formula has a very simple interpretation. At $\alpha = 0$ it correctly predicts the expected number of loops of a CM, where one pictures this exact number of triangles very "far" away from each other. When $\alpha > 0$ this finite number of triangles is multiplied by $\mathrm{e}^{6\alpha}$, giving another finite but larger amount of triangles when $N \to \infty$. In this scenario we would picture these triangles to simply be further and further apart as the system size grows. This picture will be revisited and investigated in the following section.

This approximation has been tested extensively with numerical simulations. We generated samples from (3.4) for many different degree distributions, shown in Table 3.1. The results can be seen in Figure 3.4, where we have plotted the results observed for systems of many sizes, $N \sim 100 - 4000$. In order to have a better visualization we plot values of m against a scaled parameter $\tilde{\alpha}$, given by the change of variable $\alpha = \tilde{\alpha} + \frac{1}{6} \log N$,

$$m\left(\tilde{\alpha} + \frac{1}{6} \log N \right) \approx \left(\frac{\overline{k^2}}{c} - 1 \right)^3 \mathrm{e}^{6\tilde{\alpha}} \qquad \text{for } \tilde{\alpha} \leq \tilde{\alpha_1}(N). \tag{3.18}$$

Table 3.1: List of different degree distributions used for numerical experiments done with MCMC edge swap dynamics.

type	name	formula $p(k)$	parameter values
unbounded	exponential	$exp(k\|c) = \left(\frac{c}{c+1}\right)^k \frac{1}{c+1}$	$c = 3, 4, 5, 10$
unbounded	Poissonian	$Poiss(k\|c) = \mathrm{e}^{-c}\frac{c^k}{k!}$	$c = 3, 4, 5, 10$
unbounded	power law	$PL(k) = Ak^{-\gamma} \quad k \geq 2$	$\gamma = 4$ ($\overline{k} \approx 2.5$)
bounded	bimodal	$bim(k\|3, q) = \frac{1}{2}(\delta_{k,3} + \delta_{k,q})$	$q = 5, 7, 9$
bounded	uniform	$u(k) = \frac{1}{5}\sum_{j=1}^{5}\delta_{k,j}$	-
bounded	non uniform	$v(k) = \sum_{j=1}^{5} w_j\delta_{k,j}$	$\boldsymbol{w} = (.1, .2, .3, .3, .1)$

For this regime we used waiting times of $2{\cdot}10^4$ attempted edge swaps per link (a.e.s.p.l.) and 20 samples spaced by $2{\cdot}10^3$ a.e.s.p.l. To show the accuracy of the theory with a modest number of samples, we simply plot the average of the loop density over the whole time series of loop densities between samples. We do this not only to reduce noise, but also because in fact our theory applies to the average (3.7) and not to graph instances since there is no self averaging at finite sizes. For graphs larger than 500 nodes error bars are of the order of magnitude of the markers. For smaller graphs the order magnitude can be appreciated on the right panel of Figure 3.6. In the rest of loop density plots they were omitted for clarity.

Note that this scaling in (3.18) collapses all curves of the same degree distribution up to a certain value $\tilde{\alpha_1}(N)$. As we will show in the following section the loop density at the transition is seen to go to 0 as $N \to \infty$, $m(\tilde{\alpha}_1(N)) \to 0$. This can be clearly seen in Figure 3.4.

So far, the accuracy of (3.9) suggests it could be the exact asymptotic result when $N \to \infty$. This would imply that a bias of the form (3.4) with $\alpha = \mathcal{O}(1)$ only modifies the number of expected triangles in large graphs by an $\mathcal{O}(1)$ finite amount, implying that the loop density will still vanish asymptotically. To achieve a nonvanishing loop density in the asymptotic limit a different scaling of α should be introduced, as was done in the previous chapter, [27], for 2-regular graphs, that is $p(k) = \delta_{k,2}$. Nevertheless, as will be discussed in the following section, the effect of scaling α with N is much more complicated than in the 2-regular case.

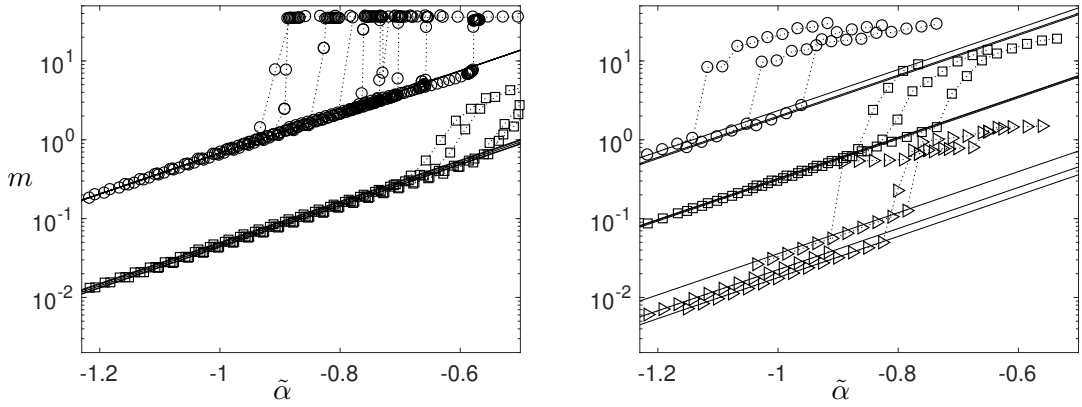


Figure 3.4: The loop density m as measured in numerical MCMC simulations of the ensemble (5.6), plotted against the rescaled control parameter $\tilde{\alpha} = \alpha - \frac{1}{6}\log N$. Left panel: $p(k) = \frac{1}{2}\delta_{k3} + \frac{1}{2}\delta_{k9}$ (circles, for system sizes $N = 100, 200, 300, 400, 500, 750, 1000, 2000, 4000$, from right to left), and $p(k) = \frac{1}{5}\sum_{j=1}^{5}\delta_{kj}$ (squares, for system sizes $N = 500, 750, 1000, 2000$, from right to left). Right panel: $p(k) = exp(k|5)$ (circles), $p(k) = Poiss(k, 5)$ (squares), and $p(k) = PL(k)$ (triangles), all for system sizes $N = 500, 1000, 2000$. See Table 3.1 for the relevant definitions. Error bars were omitted for clarity. The solid lines correspond to the corresponding theoretical prediction (3.18).

3.4 The clustered and disconnected regimes

We will now investigate the behaviour of the ensemble after the clustering transitions $\alpha_1(N)$, when (3.9) fails to reproduce the correct loop density.

For the 2-regular case the only 3-loopy structure that can exist is an isolated triangle, therefore it is possible for (3.18) to be the exact description asymptotically. For other degree distributions many other loop structures can appear in the graph. As we will now show, it seems that structures with very interacting triangles dominate entropically. Therefore the statistics of different structures need to be taken into account, making (3.9) insufficient to describe the ensemble for all values of α.

The simple picture is that when $\alpha < \alpha_1(N)$ the desired loop density is achieved by creating more triangles that are independent and far from each other; such that they do not share edges or nodes. Then, when $\alpha > \alpha_1(N)$ it is the opposite, the desired loop density is achieved by creating triangles that share edges as much as possible. The reason for this change seems to be purely entropic as the latter regime appears for all loop densities as long as the system is large enough, that is even for very small values of m. By this we mean that the transition does *not* happen because there are too many triangles and they need to start sharing edges out of lack of space, as one might guess initially. The transition

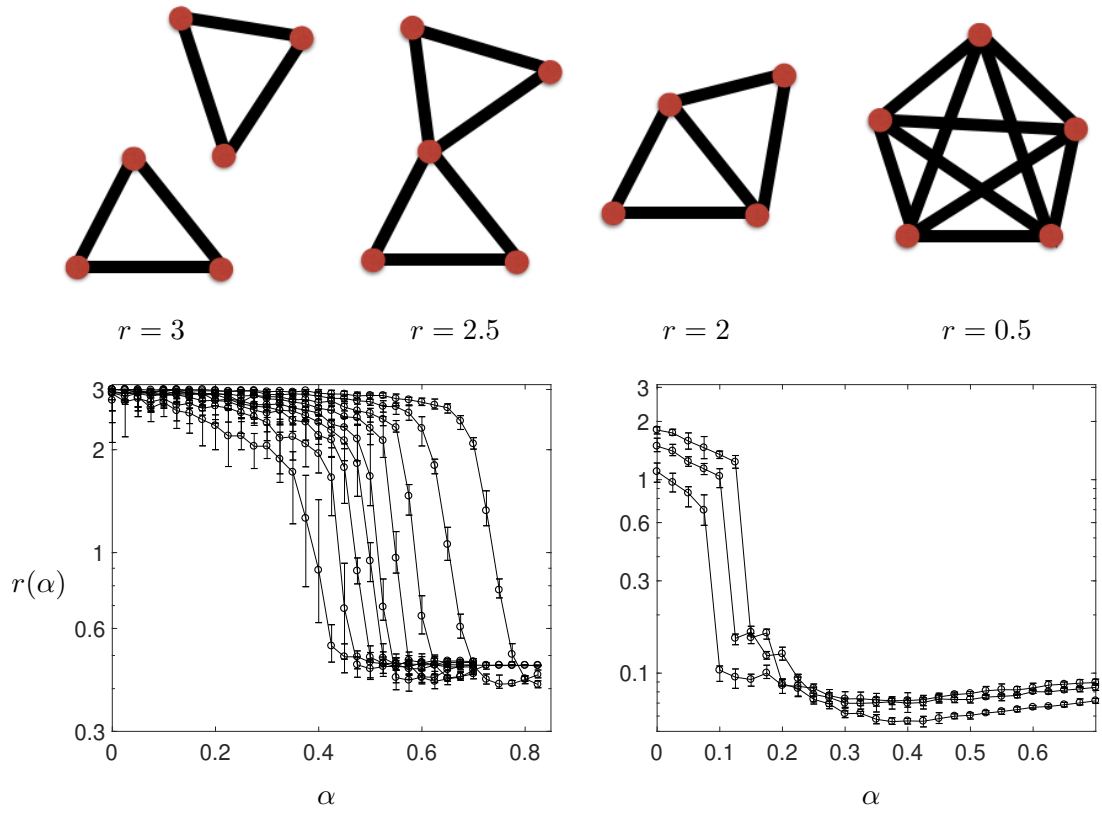


Figure 3.5: Top row: examples of small graphs and their corresponding values of $r(\mathbf{A})$. Bottom row: plots of $r(\alpha)$ as measured in simulations, shown versus α, with standard deviations shown as error bars. Left: $p(k) = \frac{1}{2}\delta_{k3} + \frac{1}{2}\delta_{5k}$, with graph sizes $N = 100, 200, 300, 400, 500, 750, 1000, 2000, 4000$ (from left to right). Right: $p(k) = \mathrm{e}^{-5}5^k/k!$, with graph sizes $N = 500, 1000, 2000$ (from left to right).

happens because for a given loop density the number of graphs one can create by "leaving triangles aside" in small clusters is larger than the number of graphs one can create by embedding them in the graph in a noninteracting way. While we have no way of proving this assertion rigorously we have a lot of numerical experiments to support this claim.

We measured the interaction between loops in samples of (3.4) using $r(\mathbf{A})$ as defined in (3.8). The empirical value of $r(\alpha) = \langle r(\mathbf{A})\rangle$ was measured in all the numerical experiments listed in Table 3.1. For values of $\alpha > \alpha_1(N)$ we increase the number of attempted edge swaps per link by factor ten, that is waiting times of $2 \cdot 10^5$ a.e.s.p.l. and $2 \cdot 10^4$ a.e.s.p.l. in between samples. In all the experiments we observe the same behavior as the two cases shown in Figure 3.5. First, an initial phase of $\langle r(\mathbf{A})\rangle \sim 3$, describing noninteracting loops, followed by a sudden drop to $\langle r(\mathbf{A})\rangle = r_{min}(N) < 1$, interacting loops. The value of α where this sudden drop occurs is what we have defined to be $\alpha_1(N)$, the graph has become clustered in order to achieve the desired loop density. This point coincides precisely with the point where (3.9) stops working, which can be seen clearly in Figure 3.1.

When increasing the system size N, it is clear that the initial part of the curves tends to flatten to a plateau at $r = 3$. This is consistent with the fact that equation (3.9), which very accurately describes the loop density in this regime, was derived assuming an underlying Poissonian distribution of triangles. The Poissonian distribution in this case was derived in [5] assuming that triangles are noninteracting, .

The remaining question is how do the other two values, $\alpha_1(N)$ and $r_{min}(N)$, behave with N.

We can distinguish the following possibilities for $r_{min}(N)$

1. $\lim_{N\to\infty} r_{min}(N) = r^* > 0$.
2. $\lim_{N\to\infty} r_{min}(N) = 0$.

Given that for a finite graph $r(\mathbf{A})$ is always lower bounded by $r = 6/(q^2 - q)$, the second option is only a possibility for unbounded graphs.

For $\alpha_1(N)$ we have the following possibilities, with the different physical implications.

1. $\lim_{N\to\infty} \alpha_1(N) = \infty$, in this case asymptotically the loop density vanishes for all values of α.
2. $\lim_{N\to\infty} \alpha_1(N) = \alpha^* > 0$. There is a first order phase transition at α^*.
3. $\lim_{N\to\infty} \alpha_1(N) = 0$. Besides $\alpha = 0$, all α values have a finite density loop density, $m(\alpha) > 0$.

We have made the distinction between bounded and unbounded distributions precisely because we believe that the behavior of these quantities might not be the same for the two cases. As can already be seen in the bound $r(\mathbf{A}) \geq 6/(q^2 - q)$, if q is growing with N, then it is possible for r to approach 0 indefinitely, contrary to the bounded case. This fact can already be appreciated in Figure 3.5, for the exponentially distributed degree distribution it is possible to see that $\langle r(\mathbf{A})\rangle$ reaches a lower value for larger N.

For the case of bounded distributions the maximum degree q asymptotically ensures there are enough nodes to create cliques that will achieve the desired loop density, see for example Figure 3.3. If the desired loop density is higher then it will just be realized with cliques of the next degree in descending order. For unbounded degree distributions this picture is not the same, since one cannot guarantee the abundance of such cliques, therefore the observed topology seems to remain connected for a longer regime in α, what we have denominated the clustered regime.

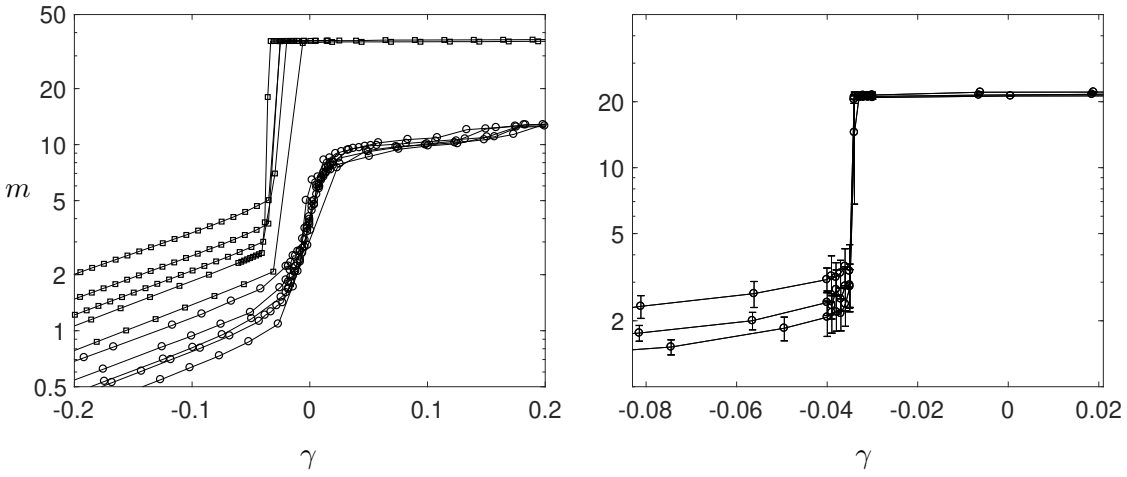


Figure 3.6: Plots of m against rescaled variable $\gamma = \alpha - [2(q+1)]^{-1} \log N$, showing the collapse of the second (shattering) transition point for different system sizes, predicted by (3.29). Left: $p(k) = \frac{1}{2}\delta_{k3} + \frac{1}{2}\delta_{k5}$ (circles), and $p(k) = \frac{1}{2}\delta_{k3} + \frac{1}{2}\delta_{k9}$ (squares). System sizes were $N = 200, 300, 400, 500, 750$, from bottom to top. Error bars are omitted to reduce cluttering. Right: close-up in the neighbourhood of the shattering transition, for $p(k) = \frac{1}{2}\delta_{k3} + \frac{1}{2}\delta_{k7}$, for system sizes $N = 200, 300, 400$ (from bottom to top). Here the error bars correspond to average plus/minus one standard deviation.

In the following subsection we show that some theory for the case bounded degree distributions can be developed.

3.4.1 Description for bounded degree distributions

As we mentioned before, numerical simulations suggest the need to include the statistics of the cliques formed by nodes of maximum degree. We denote by $K_q(\mathbf{A})$ the number of fully connected cliques of $q+1$ nodes. We now denote by $T(\mathbf{A})$ the triangles not in cliques of degree q. We can then decompose the total number of 3-loops in the following way,

$$\mathrm{Tr}(\mathbf{A}^3) = 6T(\mathbf{A}) + (q+1)q(q-1)K_q(\mathbf{A}). \tag{3.19}$$

With this decomposition we can express the generating function as follows:

$$\begin{aligned}\phi(\alpha) =& \frac{1}{N} \log \sum_{\mathbf{A}} \mathrm{e}^{6\alpha T(\mathbf{A}) + \alpha(q+1)q(q-1)K_q(\mathbf{A})} \prod_{i=1}^{N} \delta_{k_i, \sum_j A_{ij}} \\ =& \frac{1}{N} \log \sum_{T,K} P_N(T,K) \mathrm{e}^{6\alpha T} \mathrm{e}^{q(q^2-1)\alpha K} + \frac{1}{N} \log \mathcal{N}_{\mathbf{k}}.\end{aligned} \tag{3.20}$$

Where we have introduced the joint distribution of triangles and cliques for the unbiased CM.

$$\begin{aligned} P_N(T,K) =& \frac{1}{\mathcal{N}_{\mathbf{k}}} \sum_{\mathbf{A}} \delta_{T,T(\mathbf{A})} \delta_{K,K_q(\mathbf{A})} \prod_{i=1}^{N} \delta_{k_i,\sum_j A_{ij}} \\ \mathcal{N}_{\mathbf{k}} =& \sum_{\mathbf{A}} \prod_{i=1}^{N} \delta_{k_i,\sum_j A_{ij}}. \end{aligned} \tag{3.21}$$

Our main approximation is to now assume that asymptotically T, K become independent and that each is described by a Poisson distribution with expected values λ_t and $\lambda_{K_q}(N)$ respectively. This means we are again assuming that the major contribution of triangles for $T(\mathbf{A})$ comes from isolated triangles, leaving all other possible structures out of this approximation. Since isolated triangles and cliques are almost independent and rare events in the CM, one could argue that according to the *Poisson Paradigm* described by Alon and Spencer in [49], they should both be Poissonian random variables. For a similar argument regarding loops of different length see [50]. Even if this were not the exact asymptotic distribution, we make this choice to obtain a tractable solution to be tested against numerical data.

$$P_N(T,K) \sim Poiss(T|\lambda_t) Poiss(K|\lambda_{K_q}(N)). \tag{3.22}$$

We can then calculate the generating function analytically with this approximation.

$$\begin{aligned} \phi(\alpha) \approx& \frac{1}{N} \log \left\langle \mathrm{e}^{6\alpha T} \right\rangle_{Poiss(T|\lambda_t)} + \frac{1}{N} \log \left\langle \mathrm{e}^{\alpha q(q^2-1)K} \right\rangle_{Poiss(K|\lambda_{K_q})} + \frac{1}{N} \log \mathcal{N}_{\mathbf{k}} \\ =& \frac{\lambda_t}{N} \left(\mathrm{e}^{6\alpha} - 1 \right) + \frac{\lambda_{K_q}(N)}{N} \left(\mathrm{e}^{q(q^2-1)\alpha} - 1 \right) + \frac{1}{N} \log \mathcal{N}_{\mathbf{k}}. \end{aligned} \tag{3.23}$$

Which leads to the following expression for the loop density,

$$m(\alpha) = \frac{6\lambda_t}{N} \mathrm{e}^{6\alpha} + \frac{q(q^2-1)}{N} \lambda_{K_q}(N) \mathrm{e}^{q(q^2-1)\alpha}. \tag{3.24}$$

Contrary to the regular case discussed in Chapter 5, [48], there is no established rigorous result in literature for the expected number of cliques, $\lambda_K(N)$, for an arbitrary $p(k)$. Nevertheless, there is a good idea of what the correct scaling with N should be, [51]. In general for large N the expected number of appearances of isomorphisms of a given strictly balanced graph H (see [51] for definition) is expected to be $\mathcal{O}\left(N^{v(H)-e(H)}\right)$, where $e(H)$

and $v(H)$ are the number of edges and nodes of H respectively. In the case of a clique of $q+1$ nodes these are, $e(K_q) = \frac{1}{2}q(q+1)$ and $v(K_q) = q+1$. Therefore,

$$\lambda_K(N) = \mathcal{O}\left(\frac{1}{N^{\frac{1}{2}q(q-1)-1}}\right) \sim \frac{c_q}{N^{\frac{1}{2}q(q-1)-1} \cdot q(q^2-1)}. \tag{3.25}$$

We have included the factor $q(q^2-1)$ in the denominator for convenience. With this we can get the following result for small values α,

$$m(\alpha) \approx \frac{1}{N}\left(\overline{k^2}/c - 1\right)^3 \mathrm{e}^{6\alpha} + \frac{c_q}{N^{\frac{1}{2}q(q-1)}}\mathrm{e}^{\alpha q(q^2-1)}. \tag{3.26}$$

We identify in this expression two contributions. The first term corresponds to the density due to isolated triangles at low density. We denote this by $m_t(\alpha)$. The second term corresponds to the contribution of triangles present in the previously described cliques, we denote as $m_K(\alpha)$. We note that it is naturally bounded since the number of cliques of $q+1$ nodes is bounded as well. This gives,

$$m_K(\alpha) = \begin{cases} \frac{1}{N^d} c_q \mathrm{e}^{q(q^2-1)\alpha} & \text{if } \alpha \leq \alpha_2(N) \\ p(q)q(q-1) & \text{if } \alpha \geq \alpha_2(N) \end{cases}. \tag{3.27}$$

Where $d = \frac{1}{2}q(q-1)$. It is convenient to define the shattering transition as the point where all the cliques of degree q have appeared. This automatically gives us an estimate of how $\alpha_2(N)$ behaves with N. In this case we can see that it diverges logarithmically.

$$\alpha_2(N) = \frac{1}{2(q+1)}\log N + \frac{1}{q(q^2-1)}\log\left[\frac{p(q)q(q-1)}{c_q}\right]. \tag{3.28}$$

This result depends on the degree distribution $p(k)$ explicitly through q and $p(q)$, but also implicitly through c_q that is a function of the whole distribution. Note that, given that we do not know c_q, we can not test the accuracy of this prediction. For the case of regular graphs it was possible and it was very accurate as shown in Figure 5.4 Chapter 5. Nevertheless, we can do other validity tests. Equation (3.26) predicts a collapse of the curves under the following change of variable, $\alpha = \gamma + \frac{1}{2(q+1)}\log N$,

$$m\left(\gamma + \frac{1}{2(q+1)}\log N\right) \approx \begin{cases} N^{-\frac{q-2}{q+1}}\mathrm{e}^{6\gamma} & \gamma \leq \gamma_1(N) \\ c_q \mathrm{e}^{q(q^2-1)\gamma} & \text{for } \gamma_1(N) \leq \gamma \leq \gamma_2(N) \end{cases}. \tag{3.29}$$

Even though sampling very precisely in the clustering regime is very hard given that the waiting time of the MCMC algorithm is very large, overall the curves collapse nicely, as can be seen in Figure 3.6. We stress that close to the transition waiting times were so long that points on the steep part of the left panel on Figure 3.6 were probably not equilibrated for system sizes $N \geq 1000$. For this reason we show on the right panel that for system sizes $N = 200, 300, 400$ we do see an almost perfect collapse of the curves. In spite of stronger finite size effects, only for these small sizes was it possible to have more confidence in the equilibration of the MCMC algorithm so close to the transition. The predicted slope of $\frac{1}{2(q+1)}$ for $\alpha_2(N)$ in (3.28) was also tested. Results are presented in Table 3.2. We find a very good agreement for the bimodal degree distributions. For distributions $u(k)$ and $v(k)$ the prediction is close enough to the predicted value $0.8\overline{3}$, but the observed value of 0.10(1) in both cases is actually closer to what we would observe with $q = 4$. This is consistent with the fact that for this particular distributions both degrees have a similar density and $k = 4$ is more abundant in the case of $v(k)$.

With this estimate (3.28) for $\alpha_2(N)$ we can also write down an upper bound on the loop density achieved in the connected regime,

$$m_u = m_t(\alpha_2(N)) = \frac{1}{N^{\frac{q-2}{q+1}}}(\overline{k^2}/c - 1)^3 (p(q)q(q-1)/c_q)^{\frac{6}{q(q^2-1)}}. \tag{3.30}$$

This value corresponds to the loop density that would be reached if the contribution of cliques were not present. Given that cliques appear before, it is impossible to reach this density m_u in the connected phase. Even though c_q is an unknown quantity we can conclude that this number vanishes when $N \to \infty$, this is consistent with the numerical experiments, as can be seen in Figure 3.7. The results are very good for the chosen bimodal degree distributions, $p(k) = bim(k|3, q)$. In Figure 3.7 we can see the accurate prediction of two things. First, if we look at the last value of the loop density before the steep jump into the clustered phase we can see that this value actually scales with N as predicted by (3.30). Second, the end value of the jump at $\alpha_2(N)$ coincides with the prediction $p(q)q(q-1)$ as indicated by the dotted dashed line on Figure 3.7.

As a final comment, we point out that the Poissonian assumption of (3.22) implies that the shattering transition is of an entropic nature. Given that probability distributions in a uniform model as the CM are essentially proportional to the number of graphs with a

Table 3.2: Comparison of the slope of $\alpha_2(N)$ plotted against $\log N$, as measuerd from data in Figure 3.7, versus the theoretically predicted value $[2(q+1)]^{-1}$ of (3.28). The degree distributions $bim(k|a,b)$, $u(k)$ and $v(k)$ are defined as in Table 1.

$p(k)$	$bim(k\|3,5)$	$bim(k\|3,7)$	$bim(k\|3,9)$	$u(k)$	$v(k)$
theory	$0.08\bar{3}$	0.0625	0.05	$0.08\bar{3}$	$0.08\bar{3}$
simulation	0.079(5)	0.066(2)	0.057(3)	0.10(1)	0.10(1)

given property, we can study the behavior of the following quotient.

$$\frac{\mathcal{N}(\mathbf{k}|T)}{\mathcal{N}(\mathbf{k}|K)} = \frac{\#\text{ of graphs with degree sequence } \mathbf{k} \text{ and } T \text{ isolated triangles}}{\#\text{ of graphs with degree sequence } \mathbf{k} \text{ and } K \text{ q-regular cliques}}. \tag{3.31}$$

If we fixed the loop density to any arbitrary value $m^* < p(q)q(q-1)$ it can be achieved by the following number of triangles or cliques.

$$\begin{aligned} T =& \frac{m^*}{6}N, \\ K =& \frac{m^*}{q(q-1)}N. \end{aligned} \tag{3.32}$$

Using the Poissonian assumption we prove in Appendix B that

$$\lim_{N\to\infty} \frac{\mathcal{N}(\mathbf{k}|m^*N/6)}{\mathcal{N}(\mathbf{k}|m^*N/(q^2-q))} = \lim_{N\to\infty} \mathrm{e}^{-\frac{m^*}{6}\frac{q-2}{q+1}N\log N} = 0. \tag{3.33}$$

This means that no matter how small m^* is, for a large enough system there will always be more graphs that achieve it via cliques than via isolated triangles.

3.5 Discussion

In this chapter we have presented a random graph ensemble were samples are both sparse and loopy. Even though ensemble (3.4) can be regarded as the simplest random loopy graph ensemble, we can see that it has quite nontrivial behavior. While one would hope for a smooth and easy controllability of the loop density we see that in fact there are very special nontrivial regimes of α and surprisingly a very strong influence of the number of nodes in the graphs, *i.e.* the system size.

We hope that with appropriate care this ensemble could be used as a null model of loopy networks by practitioners of network science. Assume there is a real network $\mathbf{A}_0$ one wants to compare with random samples having the same loop density $m(\mathbf{A}_0)$. We

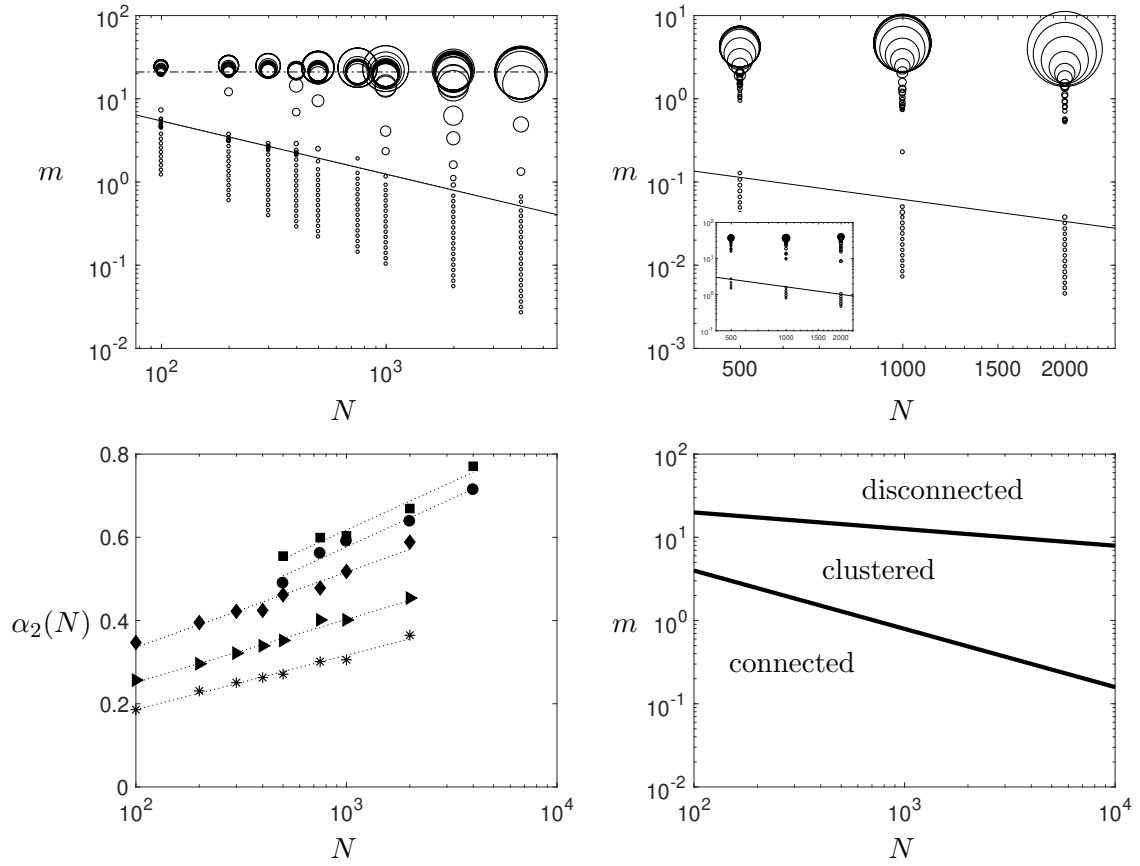


Figure 3.7: Top: scatter plots of triangle density m shown versus system size N. The width of the markers is proportional to the number of connected components. Top left: $p(k) = bim(k|3, 7)$, with dashed-dotted line corresponding to the prediction $p(q)q(q-1)$ of (3.26). The observed slope of $-0.64(5)$ is consistent with the predicted -0.625 of (3.30). Top right: $p(k) = PL(k)$, and inset $p(k) = exp(k|5)$. Here solid lines are only guides to the eye. Bottom left: linear/log plot of $\alpha_2(N)$ versus N from simulation data. Dotted line shows linear fit in good agreement with theoretical prediction (3.28), see Table 3.2. Bottom right: conjectured phase diagram of the ensemble (3.4), in the (m, N) plane.

propose that the following steps should be taken:

1. Calculate the properties of your initial graph: $\mathbf{k}(\mathbf{A}_0), m(\mathbf{A}_0), r(\mathbf{A}_0), n(\mathbf{A}_0)$.

2. Sample from (3.4) and vary α until one finds the value α^* that makes loop densities match, $m(\alpha^*) = m(\mathbf{A}_0)$. A good initial guess might be the value $\alpha_0 = \frac{1}{6}\log(m(\mathbf{A}_0)N/(\overline{k^2}/c-1)^3)$ but it is not necessary the case if $\alpha_1(N) < \alpha_0$.

3. Once loop densities are matched, compare the other properties.

   - If $n(\alpha^*) \approx n(\mathbf{A}_0)$ and $r(\alpha^*) \approx r(\mathbf{A}_0)$, then (3.4) is a suitable null model for $\mathbf{A}_0$.

- If $n(\alpha^*)$ and $n(\mathbf{A}_0)$ or $r(\alpha^*)$ and $r(\mathbf{A}_0)$ are different, it means that $\mathbf{A}_0$ is still extremely atypical in (3.4) and thus it is not a suitable null model.

Even if all observables match, it still could be the case that waiting times of the MCMC are very large. For graphs of more than a thousand nodes it could take days or more to get well mixed samples on fast multi-core computers. This just shows how the applied network scientist should be cautious when applying tools like edge swapping without a proper theory previously developed.

To summarize, we present our conjectured phase diagram in Figure 3.7 (d). With an exact solution for (3.6) one could find analytic expressions for the phase boundaries shown in Figure 3.7 (d). The main lessons are that the same loop densities may have very different topologies for different systems sizes, and that sampling anywhere outside the connected regime takes a very long time, potentially days or weeks for large graphs, even on fast multi-core machines. We conjecture that for any model any desired loop density eventually falls in the disconnected regime as N grows. For the case of bounded degree distributions with $Np(q) >> q + 1$ the clustered region practically vanishes.

There are many directions to pursue further research, ranging from practical to theoretical. From a rigorous point of view it would be interesting to see how to prove or disprove any of the assertions made in this work, that is extending rigorous results of CM beyond uniform models. Additionally, longer and more extensive simulations should be made to try to determine the exact dependence with N of $\alpha_1(N)$ and $\alpha_2(N)$, especially to find out if there is indeed a transition without scaling parameters for unbounded degree distributions.

The enormous waiting times seem to be due in part to the fact that in the clustered and disconnected phases many loops have to broken in a row to get rid of certain structures like cliques. Given that this is unlikely, an alternative MCMC with moves that involve more edges rather than only 2 could be studied, in order to speed up the algorithm and let it explore more quickly the graph space.

Finally, there are many interesting questions about the spectral properties of (3.4) to discover. First, in [48] an analytic expression for the spectral density was found for the case of regular graphs in the connected regime. The generalization to include (3.4) is presented in Chapter 6. The formation of clusters after the clustering transition points to a localization transition for the eigenvectors of $\mathbf{A}$. A similar observation has been made for dense graphs in [30] where such graphs were found to have nontrivial spectral properties,

such spectral analysis has not been done yet for the sparse case like ours.

Overall there are many open question when it comes to presenting random counter parts of real networks. It is safe to say that they are not defined by loopiness alone. It seems like real networks occupy a very small part of the abstract graph space. Finding the correct properties that will make a maximum entropy ensemble sample from a pool of realistically looking graphs is still an open problem. An alternative is to impose a constraint on the full set of eigenvalues of the adjacency matrix, in this way all loop lengths would be controlled simultaneously. This full spectral constraint has been discussed in [47, 48, 52] and will be treated in the following part of the thesis.

Part II

Maximum Entropy Loopy Graphs - Imaginary Replicas

Chapter 4

New spectral methods for random graphs and random matrices

4.1 Motivation

In the previous chapters we have given a general idea of how a degree constrained ensemble of random graphs with a bias with respect to the number of triangles behaves. Besides the fact that the number of triangles is controllable it is also safe to say that the topology is not. While an ensemble like (3.4) might be useful to model certain types of graphs with many noninteracting triangles or where triangles cluster together, it is not useful for lattice like ensembles. In Chapter 2 we have observed in numerical experiments similar behavior when one biases with respect to other geometric figures, like squares for example. The picture is very similar to the one with triangles. At first squares will appear in a noninteracting way until suddenly graphs with many disconnected fully connected bipartite graphs will appear, as shown in [6]. These are the analogue of the cliques when one biases squares instead of triangles. We already described the simplest possible case for the 2-regular graphs, and we expect a very similar picture for other degree distributions. For arbitrary degree distributions there is no easy analytic way of working with loop counters such as $n_\ell(\mathbf{A})$. Instead we need to introduce the notion of closed paths of arbitrary length ℓ. A path of length ℓ on a graph is a sequence of $\ell+1$ pairwise connected nodes. A closed path starts and ends in the same node. Loops are closed paths without repetition of nodes, except for the first and the last node. The formula for the total number of closed paths

of a given length in a graph $\mathbf{A}$ is, apart from a simple overcounting factor,

$$\mathrm{Tr}(\mathbf{A}^\ell) = \#\text{ of closed paths of length } \ell. \tag{4.1}$$

We can then write down the following model,

$$p(\mathbf{A}) = \frac{1}{Z(\boldsymbol{\alpha})} \mathrm{e}^{\sum_{\ell=3}^{K} \alpha_\ell \mathrm{Tr}(\mathbf{A}^\ell)} \prod_{i=1}^{N} \delta_{k_i, \sum_j A_{ij}}. \tag{4.2}$$

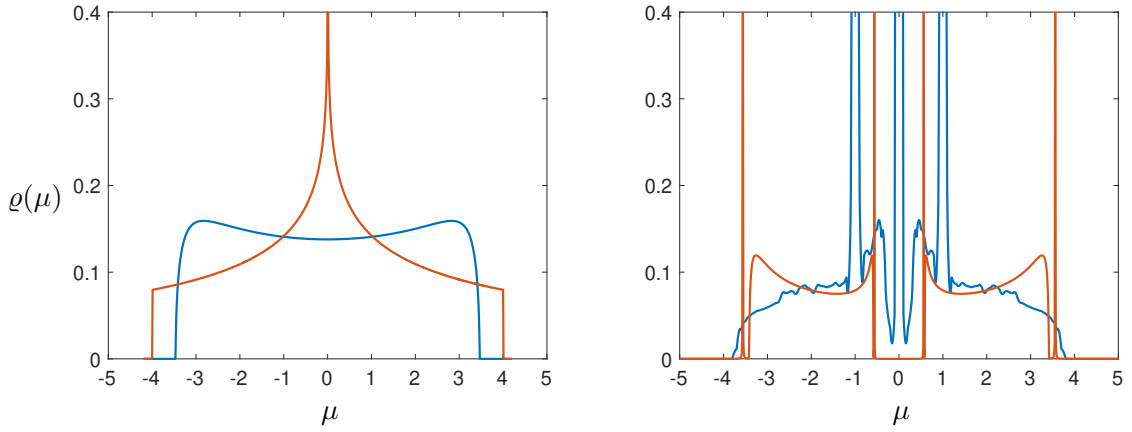
This is close to a generalization of the random graph ensemble (2.1) from Chapter 2, the main difference is that here we have a bias for each trace rather than for each loop length. That is, $\mathrm{Tr}(\mathbf{A}^\ell) = 2\ell n_\ell(\mathbf{A}) + \dots$ For example, biasing $\mathrm{Tr}(\mathbf{A}^5)$ will also favor triangles as they also contribute to that trace. This is because $\mathrm{Tr}(\mathbf{A}^\ell)$ counts the number of *all* the possible closed paths of length ℓ, including backtracking and self intersecting ones. Nevertheless, in general we can say we are biasing loops of lengths up to K. Given the phenomenology we have observed in previous sections, we can conjecture that controlling a finite number of moments will eventually lead to a graph shattering for large enough system sizes. This motivates the idea of defining an ensemble that controls all the loop lengths at once, taking $K \to \infty$.

$$p(\mathbf{A}) = \frac{1}{Z(\boldsymbol{\alpha})} \mathrm{e}^{\sum_{\ell=3}^{\infty} \alpha_\ell \mathrm{Tr}(\mathbf{A}^\ell)} \prod_{i=1}^{N} \delta_{k_i, \sum_j A_{ij}}. \tag{4.3}$$

In order to make it practically possible to assign an infinite amount of biases we use the fact that traces of matrix $\mathbf{A}$ can be expressed in terms of its eigenvalues, which in turn can be expressed as integrals over the spectral density of the matrix $\mathbf{A}$. This is only possible due to the fact we are assuming symmetric matrices.

$$\begin{aligned} \mathrm{Tr}(\mathbf{A}^\ell) =& \sum_{i=1}^{N} [\mu_i(\mathbf{A})]^\ell = N \int \mathrm{d}\mu\, \mu^\ell \varrho(\mu|\mathbf{A}), \\ \varrho(\mu|\mathbf{A}) =& \frac{1}{N} \sum_{i=1}^{N} \delta(\mu - \mu_i(\mathbf{A})). \end{aligned} \tag{4.4}$$

Given that $\mathrm{Tr}(\mathbf{A}^\ell)$ for $\ell = 0, 1, 2$ are constants equal to N, 0, $\sum_{i=1}^{N} k_i$ respectively, we can add $\alpha_{0,1,2}$ without changing the measure (4.3). Therefore, if we can find a function



$p(k) = \delta_{k,4}$ $\qquad$ $p(k) = \frac{1}{3}\delta_{k,5} + \frac{2}{3}\delta_{k,1}$

Figure 4.1: Both plots compare two spectral densities between graphs with the same degree distribution but different topological properties. The left plot shows the comparison between the asymptotic spectral distribution of a 4-regular graph (blue) and a 2 dimensional lattice (orange). The right compares a configuration model (blue) with a graph with a core-periphery structure (orange).

$\hat{\varrho}(\mu)$ such that $\hat{\varrho}(\mu) = \sum_{\ell=0}^{\infty} \alpha_\ell \mu^\ell$, then we can recast (4.3) as a functional ensemble,

$$\begin{aligned} p(\mathbf{A}) &= \frac{1}{Z[\hat{\varrho}]} \mathrm{e}^{N \int \mathrm{d}\mu \hat{\varrho}(\mu) \varrho(\mu|\mathbf{A})} \prod_{i=1}^{N} \delta_{k_i, \sum_j A_{ij}}, \\ Z[\hat{\varrho}] &= \sum_{\mathbf{A}} \mathrm{e}^{N \int \mathrm{d}\mu \hat{\varrho}(\mu) \varrho(\mu|\mathbf{A})} \prod_{i=1}^{N} \delta_{k_i, \sum_j A_{ij}}. \end{aligned} \tag{4.5}$$

This corresponds to the maximum entropy ensemble subject to a hard constraint on the degree sequence and soft constraints on the values of the spectral density at each point. For this reason we will speak of $\hat{\varrho}$ as the functional Lagrange parameter, it induces a bias for *each* value of μ.

Note that even though we motivated this definition from a topological feature such as the number of closed paths, we have arrived at an ensemble that can be interpreted by itself. That is, we can think of (4.5) as an ensemble where we are biasing certain points or regions of the spectrum to have more or fewer eigenvalues. While it is hard to picture intuitively what effect would ensue from putting more or less eigenvalues in a certain region, it clearly has important nontrivial effects. We know this by looking at the spectral distributions of certain structured matrices. For example if we compare lattices to regular graphs of the same degree, we can see that the asymptotic spectral densities are quite different, [46, 53]. All $d-$dimensional square lattices are $2d-$regular graphs. The

first difference we will see is that while the spectrum of a 2d regular graph is supported in the interval $[-2\sqrt{2d-1}, 2\sqrt{2d-1}]$, the spectrum of a square lattice will be supported in the interval $[-2d, 2d]$. The lattice spectra also seems to have a peaks at $\mu = 0$. The comparison can be seen clearly on the left panel of figure 4.1.

We find a similar situation when looking at core-periphery equitable graphs, see [41] for details on these types of graphs. In this particular case we compare graphs with only two degrees, 5 and 1. The core-periphery graph will be such that nodes are divided into two groups, core and periphery. All nodes in the core will link to exactly three other nodes in the cored and two nodes in the periphery. This gives the degree distribution $p(k) = (1/3)\delta_{k,5} + (2/3)\delta_{k,1}$. In the right panel of Figure 4.1 we compare the spectral distribution of the core-periphery graph with that of a configuration model with the same degree distribution. There is a very noticeable and clear difference, there is a massive gap in the middle of the spectral density of the core-periphery graph.

In summary we can say that,

- The lattice structure "pushed" the eigenvalues outside the interval $[-2\sqrt{2d-1}, 2\sqrt{2d-1}]$.
- The core-periphery structure created a gap in the neighborhood of $\mu = 0$.

Given the above observations, it is natural to try to work in the inverse way: move eigenvalues around to create interesting topological structures. While we will not pursue this idea further in the thesis we do believe that it is an important motivation for the study of (4.5). In [34] other loopy structures were explored showing similar effects of the loopiness on the spectral density.

We have then arrived with the choice (4.5) at a very flexible ensemble. We can recover simpler models by simply choosing $\hat{\varrho}(\mu) = \alpha\mu^3$, that would recover (3.4). We can choose biasing just certain specific eigenvalue regions, for example we could choose just biasing the number of eigenvalues in the interval $[a, b]$ $\hat{\varrho}(\mu) = y[\theta(\mu - a) - \theta(\mu - b)]$. This would lead to equations similar to the ones developed in other works looking at large deviations of eigenvalues inside a given interval, [54–58].

Dealing with an ensemble like (4.5) is clearly not a trivial analytical task. We will present a possible approach based on the method of imaginary replicas. Although it might not be applicable for all possible choices of $\hat{\varrho}$, it is certainly useful in many instances both for graphs and for random matrices in general. We will show how the full formalism works

in the following chapters for unweighted graphs with arbitrary degree distributions. For weighted graphs simpler versions of this approach that also use imaginary replicas have been successfully used in [54–58].

4.2 Introducing a functional approach for spectrally constrained ensembles

To keep this discussion more general we will introduce the following family of tilted ensembles for random matrices $\boldsymbol{H}$. We will change notation from $\mathbf{A}$ to $\boldsymbol{H}$ to stress that this discussion could be applied to any sort of random matrix ensemble including weighted graphs and fully connected ensembles. Suppose we have a matrix ensemble $p_0(\boldsymbol{H})$, we can then define the following tilted ensemble.

$$\begin{aligned} p(\boldsymbol{H}) =& \frac{1}{Z[\hat{\varrho}]} \mathrm{e}^{N\int \mathrm{d}\mu \hat{\varrho}(\mu)\varrho(\mu|\boldsymbol{H})} p_0(\boldsymbol{H}), \\ Z[\hat{\varrho}] =& \left\langle \mathrm{e}^{N\int \mathrm{d}\mu \hat{\varrho}(\mu)\varrho(\mu|\boldsymbol{H})} \right\rangle_{p_0(\boldsymbol{H})}. \end{aligned} \tag{4.6}$$

The objective is to calculate the following generating function,

$$\begin{aligned} \phi[\hat{\varrho}] =& \frac{1}{N} \log \left\langle \mathrm{e}^{N\int \mathrm{d}\mu \hat{\varrho}(\mu)\varrho(\mu|\boldsymbol{H})} \right\rangle_0, \\ \frac{\delta\phi}{\delta\hat{\varrho}(\mu)} =& \langle \varrho(\mu|\boldsymbol{H})\rangle_{p(\boldsymbol{H})}. \end{aligned} \tag{4.7}$$

Where we use the notation $\langle\rangle_0 := \langle\rangle_{p_0(\boldsymbol{H})}$

This theory is functional in nature as now the theory depends on the function $\hat{\varrho}$. The derivative is defined as a functional derivative.

4.2.1 Transforming the ensemble into a suitable form with the Stieltjes transform

We need to introduce a way to deal with the integral analytically, so we represent it via the Sokhotski–Plemelj formula, [35], for $a < 0 < b$

$$\lim_{\varepsilon\to 0} \int_a^b \mathrm{d}\mu \frac{f(\mu)}{\mu - \mathrm{i}\varepsilon} = \mathrm{i}\pi f(0) + \mathcal{P}\int_a^b \mathrm{d}\mu \frac{f(\mu)}{\mu} \tag{4.8}$$

Where $\mathcal{P}$ corresponds to the Cauchy principal value.

With this result we can rewrite our integral in the following way,

$$N\int \mathrm{d}\mu\hat{\varrho}(\mu)\varrho(\mu|\boldsymbol{H}) = \sum_{i=1}^{N}\hat{\varrho}(\mu_i) = \lim_{\varepsilon\to 0}\sum_{i=1}^{N}\frac{1}{\pi}\mathrm{Im}\int_{-\mu_0}^{\mu_0}\mathrm{d}\mu\hat{\varrho}(\mu)\frac{1}{\mu_i-\mu-\mathrm{i}\varepsilon}$$
$$= \lim_{\varepsilon\to 0}\int_{-\mu_0}^{\mu_0}\mathrm{d}\mu\hat{\varrho}(\mu)\frac{1}{\pi}\mathrm{Im}\mathrm{Tr}(\boldsymbol{H}-\mu_\varepsilon\mathbf{I})^{-1} \tag{4.9}$$

Which means that we are effectively approximating the true spectral density with the $\varepsilon-$smooth version of the spectral density (1.28), discussed previously in Chapter 1. Note that while the integral would diverge if taken over the whole real line, the bound μ_0 keeps it convergent. We can then consider that at finite ε we are working with an ε ensemble, with generating function given by

$$\phi_\varepsilon[\hat{\varrho}] = \frac{1}{N}\log\left\langle \exp\left(\int_{-\mu_0}^{\mu_0}\mathrm{d}\mu\hat{\varrho}(\mu)\frac{1}{\pi}\mathrm{Im}\mathrm{Tr}(\boldsymbol{H}-\mu_\varepsilon\mathbf{I})^{-1}\right)\right\rangle_0,$$
$$\frac{\delta\phi_\varepsilon}{\delta\hat{\varrho}(\mu)} = \mathrm{Im}\left\langle\frac{1}{N\pi}\mathrm{Tr}(\boldsymbol{H}-\mu_\varepsilon\mathbf{I})^{-1}\right\rangle_{p(\boldsymbol{H})}. \tag{4.10}$$

This formulation allows us to highlight that at finite $\varepsilon > 0$ the functional derivative must be the imaginary part of an analytic function, a harmonic function, at least when the average over $p(\boldsymbol{H})$ is over a finite set (for example in the case of unweighted graphs). In order to continue with the calculation we then need to work with the integral. For that we will carry out the following manipulations in order to be able to approximate the eigenvalue integral with a Riemann sum.

$$N\int \mathrm{d}\mu\hat{\varrho}(\mu)\varrho(\mu|\boldsymbol{H}) = \sum_{i=1}^{N}\hat{\varrho}(\mu_i) = \lim_{\varepsilon\to 0}\sum_{i=1}^{N}\frac{1}{\pi}\mathrm{Im}\int_{-\mu_0}^{\mu_0}\mathrm{d}\mu\hat{\varrho}(\mu)\frac{1}{\mu_i-\mu-\mathrm{i}\varepsilon}$$
$$= \lim_{\varepsilon\to 0}\int_{-\mu_0}^{\mu_0}\mathrm{d}\mu\hat{\varrho}(\mu)\left[-\frac{1}{\pi}\mathrm{Im}\sum_{i=1}^{N}\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Log}(\mu_i-\mu_\varepsilon)\right] = \lim_{\varepsilon\to 0}\int_{-\mu_0}^{\mu_0}\mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathcal{I}_\varepsilon(\mu|\boldsymbol{H})$$
$$= \hat{\varrho}(\mu_0)\mathcal{I}(\mu_0|\boldsymbol{H}) - \hat{\varrho}(-\mu_0)\mathcal{I}(-\mu_0|\boldsymbol{H}) - \lim_{\varepsilon\to 0}\int_{-\mu_0}^{\mu_0}\mathrm{d}\mu\hat{\varrho}'(\mu)\mathcal{I}_\varepsilon(\mu|\boldsymbol{H}). \tag{4.11}$$

Where we have introduced an ε smoothened counting function $\mathcal{I}_\varepsilon(\mu|\boldsymbol{H})$, such that

$$\mathcal{I}_\varepsilon(\mu|\boldsymbol{H}) = \sum_{i=1}^{N} -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Log}(\mu_i-\mu_\varepsilon) \xrightarrow{\varepsilon\to 0} \mathcal{I}(\mu|\boldsymbol{H}) = \sum_{i=1}^{N}\theta(\mu-\mu_i). \tag{4.12}$$

Where

$$\theta(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{2} & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases} . \tag{4.13}$$

Note that we have assumed that we can choose an interval $[-\mu_0, \mu_0]$ to contain all eigenvalues. This can always be done for a fixed $\boldsymbol{H}$, that is $\mu_0 = \mu_0(\boldsymbol{H})$ such that $\mathcal{I}(\mu_0|\boldsymbol{H}) = N$ and $\mathcal{I}(-\mu_0|\boldsymbol{H}) = 0$.

The final transformation is to rotate the argument in the following way,

$$-\frac{1}{\pi}\mathrm{Im}\,\mathrm{Log}(\mu_i - \mu_\varepsilon) = -\frac{1}{\pi}\mathrm{Im}\,\mathrm{Log}[\mathrm{i}(\mu_i - \mu_\varepsilon)] + \frac{1}{2}. \tag{4.14}$$

Where

$$-\frac{1}{\pi}\mathrm{Im}\,\mathrm{Log}[\mathrm{i}(\mu_i - \mu_\varepsilon)] = -\frac{1}{\pi}\arctan\left(\frac{\mu_i - \mu}{\varepsilon}\right) \xrightarrow{\varepsilon\to 0} \frac{1}{2}\mathrm{sgn}(\mu - \mu_i)$$

$$\mathrm{sgn}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases} . \tag{4.15}$$

We finally write this as,

$$\begin{aligned} N\int \mathrm{d}\mu\hat{\varrho}(\mu)\varrho(\mu|\boldsymbol{H}) =& \frac{N}{2}(\hat{\varrho}(\mu_0) + \hat{\varrho}(-\mu_0)) - \lim_{\varepsilon\to 0}\int_{-\mu_0}^{\mu_0}\mathrm{d}\mu\hat{\varrho}'(\mu)\frac{2}{\pi}\mathrm{Im}\sum_{i=1}^{N}\frac{1}{2}\mathrm{Log}\left[\frac{2\pi}{\mathrm{i}(\mu_i - \mu_\varepsilon)}\right] \\ =& \frac{N}{2}(\hat{\varrho}(\mu_0) + \hat{\varrho}(-\mu_0)) - \lim_{\varepsilon\to 0}\int_{-\mu_0}^{\mu_0}\mathrm{d}\mu\hat{\varrho}'(\mu)\mathrm{Im}\frac{2}{\pi}\log\prod_{i=1}^{N}\sqrt{\frac{2\pi}{\mathrm{i}(\mu_i - \mu_\varepsilon)}}. \end{aligned} \tag{4.16}$$

Note that we have been using the notation Log to represent the principal branch of logarithm, but in this case we are changing notation to $\log f(z)$ to represent the logarithm of a complex function. In general it is defined by

$$\log f(z) \stackrel{\mathrm{def}}{:=} b + \int_a^z \mathrm{d}\xi\frac{f'(\xi)}{f(\xi)}. \tag{4.17}$$

Where one has to specify the integration path, and the complex constants b and a. Note that if f'/f has no poles then it will be independent of the chosen path. In our case f'/f

is exactly what we want,

$$\frac{\frac{\mathrm{d}}{\mathrm{d}z}\prod_{i=1}^{N}\sqrt{\frac{2\pi}{\mathrm{i}(\mu_i-z)}}}{\prod_{i=1}^{N}\sqrt{\frac{2\pi}{\mathrm{i}(\mu_i-z)}}}=\frac{1}{2}\sum_{i=1}^{N}\frac{1}{\mu_i-z}=\frac{1}{2}\mathrm{Tr}(\boldsymbol{H}-z\mathbf{I})^{-1}. \tag{4.18}$$

Note that we can express all of this in terms of a complex Gaussian integral.

$$Z(z|\boldsymbol{H})=\int_{\mathbb{R}^N}\mathrm{d}\boldsymbol{\phi}\,\mathrm{e}^{-\frac{\mathrm{i}}{2}\boldsymbol{\phi}\cdot(\boldsymbol{H}-z\mathbf{I})\boldsymbol{\phi}}=\prod_{i=1}^{N}\sqrt{\frac{2\pi}{\mathrm{i}(\mu_i-z)}} \tag{4.19}$$

Where this is only well defined when $\mathrm{Im}z>0$. Therefore we can define the logarithm without ambiguity for z in the upper half plane and write for the *log* of $Z(z|\boldsymbol{H})$:

$$\begin{aligned}\log\prod_{i=1}^{N}\sqrt{\frac{2\pi}{\mathrm{i}(\mu_i-z)}}=\log Z(z|\boldsymbol{H})=&\sum_{i=1}^{N}\frac{1}{2}\mathrm{Log}\left[\frac{2\pi}{\mathrm{i}(\mu_i-a)}\right]+\int_a^z\mathrm{d}\xi\frac{1}{2}\mathrm{Tr}(\boldsymbol{H}-\xi\mathbf{I})^{-1}\\=&\sum_{i=1}^{N}\frac{1}{2}\mathrm{Log}\left[\frac{2\pi}{\mathrm{i}(\mu_i-z)}\right].\end{aligned} \tag{4.20}$$

We get to our final expression for the integral,

$$\begin{aligned}N\int\mathrm{d}\mu\hat{\varrho}(\mu)\varrho(\mu|\boldsymbol{H})=&\frac{N}{2}(\varrho(\mu_0)+\varrho(-\mu_0))+\lim_{\varepsilon\to0}\int_{-\mu_0}^{\mu_0}\mathrm{d}\mu\hat{\varrho}'(\mu)\left[-\frac{2}{\pi}\mathrm{Im}\log Z(\mu_\varepsilon|\boldsymbol{H})\right]\\=&\frac{N}{2}(\hat{\varrho}(\mu_0)+\hat{\varrho}(-\mu_0))+\lim_{\varepsilon\to0}\lim_{\Delta\to0}\sum_{\mu}\frac{\Delta}{\pi}\hat{\varrho}'(\mu)\left[-2\mathrm{Im}\log Z(\mu_\varepsilon|\boldsymbol{H})\right].\end{aligned} \tag{4.21}$$

Given that we have now mapped it to a Riemann integral we can approximate it by a Riemann sum. We have simplified the notation in the following way,

$$\lim_{\Delta\to0}\sum_{\mu}\Delta f(\mu):=\lim_{M\to\infty}\sum_{i=1}^{M}\Delta f(\mu_i)\qquad\mu_i=-\mu_0+(i-1)\frac{2\mu_0}{M-1}\qquad\Delta=\frac{2\mu_0}{M-1}. \tag{4.22}$$

We have intentionally made the effort to write the integral as a limit of ε and Δ to justify the following approach. We can now define a family of ε,Δ distributions such that

$$\begin{aligned}p_{\varepsilon,\Delta}(\boldsymbol{H})=&\frac{1}{Z_{\varepsilon,\Delta}[\hat{\varrho}]}\exp\left(\frac{N}{2}(\hat{\varrho}(\mu_0)+\hat{\varrho}(-\mu_0))+\sum_{\mu}\frac{\Delta}{\pi}\hat{\varrho}'(\mu)\left[-2\mathrm{Im}\log Z(\mu_\varepsilon|\boldsymbol{H})\right]\right)p_0(\boldsymbol{H}),\\\phi_{\varepsilon,\Delta}[\hat{\varrho}]=&\frac{1}{N}\log\left\langle\mathrm{e}^{\frac{N}{2}(\hat{\varrho}(\mu_0)+\hat{\varrho}(-\mu_0))+\sum_{\mu}\frac{\delta}{\pi}\hat{\varrho}'(\mu)[-2\mathrm{Im}\log Z(\mu_\varepsilon|\boldsymbol{H})]}\right\rangle.\end{aligned} \tag{4.23}$$

Given that we have the identity in (4.21), if we assume that the support of $p_0(\boldsymbol{H})$ is finite, as in the case of unweighted simple graphs, for a fixed N we can define a global pair ε, Δ such that

$$\forall \boldsymbol{H} : \left| \int \mathrm{d}\mu \hat{\varrho}(\mu) \varrho(\mu|\boldsymbol{H}) - \frac{1}{2}(\hat{\varrho}(\mu_0) + \hat{\varrho}(-\mu_0)) + \sum_{\mu} \Delta \hat{\varrho}'(\mu) \frac{2}{\pi} \mathrm{Im} \log Z(\mu_\varepsilon|\boldsymbol{H}) \right| < \delta. \quad (4.24)$$

It is only a matter of fixing $\mu_0 = \max_{\boldsymbol{H}} \mu_0(\boldsymbol{H})$, and then $(\varepsilon, \Delta) = \min_{\boldsymbol{H}}(\varepsilon(\boldsymbol{H}), \Delta(\boldsymbol{H}))$.

The existence of these global parameters allows us to be able to confidently approximate the generating function of the original ensemble with the generating function of the ε, Δ ensemble and afterwards taking the limit $\varepsilon, \Delta \to 0$

$$\phi[\hat{\varrho}] = \lim_{\varepsilon \to 0} \lim_{\Delta \to 0} \frac{1}{N} \log \sum_{\boldsymbol{H}} p_0(\boldsymbol{H}) \mathrm{e}^{\frac{N}{2}(\hat{\varrho}(\mu_0) + \hat{\varrho}(-\mu_0)) + \sum_{\mu} \frac{\Delta}{\pi} \hat{\varrho}'(\mu)[-2\mathrm{Im} \log Z(\mu_\varepsilon|\boldsymbol{H})]}. \quad (4.25)$$

Finally we can rewrite this using the identity $\mathrm{Im}\, z = -\frac{\mathrm{i}}{2}(z - \overline{z})$as

$$\phi[\hat{\varrho}] = \lim_{\varepsilon, \Delta \to 0} \frac{1}{N} \log \left\langle \prod_{\mu} \exp\left(\mathrm{i} \frac{\Delta}{\pi} \hat{\varrho}'(\mu)[\log Z(\mu_\varepsilon|\boldsymbol{H})] - \mathrm{i} \frac{\Delta}{\pi} \hat{\varrho}'(\mu)[\log \overline{Z(\mu_\varepsilon|\boldsymbol{H})}] \right) \right\rangle. \quad (4.26)$$

4.2.2 The imaginary replica trick

Note that so far we have taken extreme care in only using expressions with complex numbers that are precisely defined, our definition of log (4.20) and $\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$. Given ε, Δ we can regard the generating function as a function of $2M$ complex numbers,

$$\begin{aligned} \phi_{\varepsilon,\Delta} :& \mathbb{C}^M \times \mathbb{C}^M \to \mathbb{C} \\ \phi_{\varepsilon,\Delta}[\{w_\mu, \tau_\mu\}_\mu] =& \frac{1}{N} \log \left\langle \prod_{\mu} \exp\left(w_\mu [\log Z(\mu_\varepsilon|\boldsymbol{H})] + \tau_\mu [\log \overline{Z(\mu_\varepsilon|\boldsymbol{H})}] \right) \right\rangle. \end{aligned} \quad (4.27)$$

Where the chosen branch for log is such that,

$$\frac{\partial}{\partial w_{\mu'}} \phi_{\varepsilon,\Delta} = \frac{1}{N} \frac{\left\langle \prod_{\mu} \exp\left(w_\mu [\log Z(\mu_\varepsilon|\boldsymbol{H})] + \tau_\mu [\log \overline{Z(\mu_\varepsilon|\boldsymbol{H})}] \right) \cdot [\log Z(\mu_\varepsilon{}'|\boldsymbol{H})] \right\rangle_0}{\left\langle \prod_{\mu} \exp\left(w_\mu [\log Z(\mu_\varepsilon|\boldsymbol{H})] + \tau_\mu [\log \overline{Z(\mu_\varepsilon|\boldsymbol{H})}] \right) \right\rangle_0}. \quad (4.28)$$

This means that we want $\phi_{\varepsilon,\Delta}$ to be the cumulant generating function (divided by N) for the random vector $(\log Z(\mu_\varepsilon|\boldsymbol{H}), \log \overline{Z(\mu_\varepsilon|\boldsymbol{H})})_\mu$. It is a multivariate power series where the coefficients can be obtained by noting that taking derivatives of $\phi_{\varepsilon,\Delta}$ and then setting

all the arguments to 0 should always give the cumulants of the corresponding random object, for example,

$$\frac{\partial^4}{\partial w_\mu \partial \tau_{\mu'} \partial^2 w_{\mu''}} \phi_{\varepsilon,\Delta}|_{\{w_\mu,\tau_\mu\}\to 0} = \frac{1}{N}\left\langle \log Z(\mu_\varepsilon|\boldsymbol{H}) \log \overline{Z(\mu_\varepsilon{}'|\boldsymbol{H})} [\log Z(\mu_\varepsilon{}''|\boldsymbol{H})]^2 \right\rangle_{c,0}. \tag{4.29}$$

We will now proceed with the replica tick which consists in finding the value of this function just at nonnegative integers. In this way we can perform the average over $p_0(\boldsymbol{H})$,

$$\begin{aligned}
\phi_{\varepsilon,\Delta}[\{n_\mu, m_\mu\}_\mu] =& \frac{1}{N} \log \left\langle \prod_\mu Z(\mu_\varepsilon|\boldsymbol{H})^{n_\mu} \overline{Z(\mu_\varepsilon|\boldsymbol{H})}^{m_\mu} \right\rangle_0 \\
=& \frac{1}{N} \log \left\langle \prod_\mu \left(\int \mathrm{d}\boldsymbol{\phi}\, \mathrm{e}^{-\frac{\mathrm{i}}{2}\boldsymbol{\phi}\cdot(\boldsymbol{H}-\mu_\varepsilon \mathbf{I})\boldsymbol{\phi}} \right)^{n_\mu} \overline{\left(\int \mathrm{d}\boldsymbol{\psi}\, \mathrm{e}^{-\frac{\mathrm{i}}{2}\boldsymbol{\psi}\cdot(\boldsymbol{H}-\mu_\varepsilon \mathbf{I})\boldsymbol{\psi}} \right)}^{m_\mu} \right\rangle_0 \\
=& \frac{1}{N} \log \int \prod_{i=1}^N \prod_\mu \prod_{\alpha_\mu=1}^{n_\mu} \mathrm{d}\phi^i_{\mu,\alpha_\mu} \prod_{\beta_\mu=1}^{m_\mu} \mathrm{d}\psi^i_{\mu,\beta_\mu} \\
\times \left\langle \exp\left(-\frac{\mathrm{i}}{2} \sum_{\mu,\alpha_\mu} \sum_{ij} \phi^i_{\mu,\alpha_\mu} \phi^j_{\mu,\alpha_\mu} (H_{ij} - \mu_\varepsilon \delta_{ij}) + \frac{\mathrm{i}}{2} \sum_{\mu,\beta_\mu} \sum_{ij} \psi^i_{\mu,\beta_\mu} \psi^j_{\mu,\beta_\mu} (H_{ij} - \overline{\mu_\varepsilon} \delta_{ij}) \right) \right\rangle_0 .
\end{aligned} \tag{4.30}$$

Introducing the following $d-$dimensional vector with $d = \sum_\mu n_\mu + m_\mu$,

$$\boldsymbol{\varphi} = \begin{pmatrix} \{\phi_{\mu,\alpha_\mu}\} \\ \{\psi_{\mu,\beta_\mu}\} \end{pmatrix}, \tag{4.31}$$

and the dot product

$$\boldsymbol{\varphi} \cdot \boldsymbol{\varphi}' = \sum_{\mu,\alpha_\mu} \phi_{\mu,\alpha_\mu} \phi'_{\mu,\alpha_\mu} - \sum_{\mu,\beta_\mu} \psi_{\mu,\beta_\mu} \psi'_{\mu,\beta_\mu}. \tag{4.32}$$

We get the following simplified expression,

$$\begin{aligned}
\phi_{\varepsilon,\Delta}[\{n_\mu, m_\mu\}] =& \frac{1}{N} \log \int \prod_{i=1}^N \mathrm{d}\boldsymbol{\varphi}^i \mathrm{e}^{\frac{\mathrm{i}}{2}\sum_{i=1}^N \boldsymbol{\varphi}^i \cdot \boldsymbol{\mu} \boldsymbol{\varphi}^i} \left\langle \mathrm{e}^{-\frac{\mathrm{i}}{2}\sum_{ij} \boldsymbol{\varphi}^i \cdot \boldsymbol{\varphi}^j H_{ij}} \right\rangle_0 \\
=& \frac{1}{N} \log \int \prod_{i=1}^N \mathrm{d}\boldsymbol{\varphi}^i \, \exp\left(\frac{\mathrm{i}}{2} \sum_{i=1}^N \boldsymbol{\varphi}^i \cdot \boldsymbol{\mu} \boldsymbol{\varphi}^i + N \chi[\{\boldsymbol{\varphi}^i \cdot \boldsymbol{\varphi}^j\}_{i,j=1,\ldots,N}] \right) \\
\chi[\mathbf{B}] =& \frac{1}{N} \log \left\langle \mathrm{e}^{-\frac{\mathrm{i}}{2}\mathrm{Tr}(\mathbf{B}\boldsymbol{H})} \right\rangle_0 .
\end{aligned} \tag{4.33}$$

Our final remark regards the log in the exponent. While a log present in the exponent is not requiring the branch to be specified, it is necessary to choose one that will be suitable for calculations. Therefore we chose log such that the resulting function is the cumulant generating function for the ensemble $p_0(\boldsymbol{H})$, that is, it is defined by the relation

$$\frac{\partial}{\partial B_{ij}} \log \left\langle \mathrm{e}^{-\frac{\mathrm{i}}{2}\mathrm{Tr}(\mathbf{B}\boldsymbol{H})} \right\rangle_0 = \frac{\left\langle H_{ij} \mathrm{e}^{-\frac{\mathrm{i}}{2}\mathrm{Tr}(\mathbf{B}\boldsymbol{H})} \right\rangle_0}{\left\langle \mathrm{e}^{-\frac{\mathrm{i}}{2}\mathrm{Tr}(\mathbf{B}\boldsymbol{H})} \right\rangle_0}. \tag{4.34}$$

Since the function of which log is taken has no zeros there should not be any problem with choosing any integration path. Our final result is then the following identity,

$$\begin{aligned} \phi[\hat{\varrho}] &= \lim \frac{1}{N} \log \int \prod_{i=1}^{N} \mathrm{d}\boldsymbol{\varphi}^i \, \exp\left(\frac{\mathrm{i}}{2} \sum_{i=1}^{N} \boldsymbol{\varphi}^i \cdot \boldsymbol{\mu} \boldsymbol{\varphi}^i + N \chi[\{\boldsymbol{\varphi}^i \cdot \boldsymbol{\varphi}^j\}_{i,j=1,\ldots,N}] \right), \\ \lim &= \lim_{\varepsilon \to 0} \lim_{\Delta \to 0} \lim_{n_\mu \to \mathrm{i}\frac{\Delta}{\pi}\hat{\varrho}'(\mu)} \lim_{m_\mu \to -n_\mu}. \end{aligned} \tag{4.35}$$

While compact, this is still a very complicated expression, due to the presence of the cumulant generating function. At this point one would want to proceed with the integral through a saddle point argument. However, it is not strictly possible to do so since formally the dimensionality should not grow with N, and in this case we are integrating over $\mathbb{R}^{d\times N}$. Therefore it is necessary to reduce or at least fix the dimensionality in such a way that N just becomes a parameter.

To achieve our goal we need to exploit the fact that we will be mostly working with ensembles that do not privilege any particular node, this means that a probability of a given matrix is invariant under permutations. This means that any appearance of $\boldsymbol{\varphi}^i$ will be preceded by a $\sum_i$, therefore it should be possible to write everything in terms of the distribution,

$$P(\boldsymbol{\varphi}) = \frac{1}{N} \sum_{i=1}^{N} \delta(\boldsymbol{\varphi} - \boldsymbol{\varphi}^i). \tag{4.36}$$

We are assuming then that we can write the cumulant generating function in the following

way,

$$\begin{aligned}\chi[\{\boldsymbol{\varphi}^i\cdot\boldsymbol{\varphi}^j\}_{i,j}] =& \sum_i \kappa_1(\boldsymbol{\varphi}^i,N)+\sum_{i,j}\kappa_2(\boldsymbol{\varphi}^i,\boldsymbol{\varphi}^j,N)+\sum_{i,j,k}\kappa_3(\boldsymbol{\varphi}^i,\boldsymbol{\varphi}^j,\boldsymbol{\varphi}^k,N)+\dots\\ =&N\int \mathrm{d}\boldsymbol{\varphi}P(\boldsymbol{\varphi})\kappa_1(\boldsymbol{\varphi},N)+N^2\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'P(\boldsymbol{\varphi})P(\boldsymbol{\varphi}')\kappa_2(\boldsymbol{\varphi},\boldsymbol{\varphi}',N)\\ &+N^3\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\mathrm{d}\boldsymbol{\varphi}''P(\boldsymbol{\varphi})P(\boldsymbol{\varphi}')P(\boldsymbol{\varphi}'')\kappa_3(\boldsymbol{\varphi},\boldsymbol{\varphi}',\boldsymbol{\varphi}'',N)+\dots \quad . \end{aligned} \tag{4.37}$$

We then proceed in the typical way by enforcing the definition (4.36) through the use of delta distributions. In this case functional delta distributions are needed, which we introduce by discretizing $\mathbb{R}^d$ into hypercubes of hypervolume $\Delta_{\boldsymbol{\varphi}}$ and introducing a Dirac delta at each point. The conjugate parameter to $P(\boldsymbol{\varphi})$ is scaled appropriately.

$$\begin{aligned}1 &= \prod_{\boldsymbol{\varphi}}\left[\int_{-\infty}^{\infty}\frac{\mathrm{d}P(\boldsymbol{\varphi})\mathrm{d}\hat{P}(\boldsymbol{\varphi})}{2\pi}\mathrm{e}^{\mathrm{i}\hat{P}(\boldsymbol{\varphi})(P(\boldsymbol{\varphi})-\frac{1}{N}\sum_{i=1}^N\delta(\boldsymbol{\varphi}-\boldsymbol{\varphi}^i))}\right]\\ &= \int\prod_{\boldsymbol{\varphi}}\left[\frac{\mathrm{d}P(\boldsymbol{\varphi})\mathrm{d}\hat{P}(\boldsymbol{\varphi})}{2\pi/(\Delta_{\boldsymbol{\varphi}}N)}\right]\mathrm{e}^{N\mathrm{i}\sum_{\boldsymbol{\varphi}}\Delta_{\boldsymbol{\varphi}}\hat{P}(\boldsymbol{\varphi})[P(\boldsymbol{\varphi})-\sum_{i=1}^N\delta(\boldsymbol{\varphi}-\boldsymbol{\varphi}^i)]}\\ 1 &\xrightarrow{\Delta_{\boldsymbol{\varphi}}\to 0}\int\mathcal{D}P\mathcal{D}\hat{P}\mathrm{e}^{N\mathrm{i}\int\mathrm{d}\boldsymbol{\varphi}P(\boldsymbol{\varphi})\hat{P}(\boldsymbol{\varphi})-\mathrm{i}\sum_{i=1}^N\hat{P}(\boldsymbol{\varphi}^i)}\end{aligned} \tag{4.38}$$

By taking the limit $\Delta_{\boldsymbol{\varphi}}\to 0$ the above integral has formally become ill-defined, but this is typical for field integrals of this kind. It is important to have this in mind, but ultimately we are not constructing a rigorous mathematical proof but rather a way of calculating observables that we will validate via numerical sampling of the ensemble. It might also lead to correct results that can be proven later once the answer is known.

We finally arrive to the general result,

$$\boldsymbol{\varphi}[\hat{\varrho}] = \lim\frac{1}{N}\log\int\mathcal{D}P\mathcal{D}\hat{P}\,\mathrm{e}^{N\mathrm{i}\int\mathrm{d}\boldsymbol{\varphi}P(\boldsymbol{\varphi})\hat{P}(\boldsymbol{\varphi})+N\chi^*[P]+N\log\int\mathrm{d}\boldsymbol{\varphi}\,\mathrm{e}^{\frac{\mathrm{i}}{2}\boldsymbol{\varphi}\cdot\boldsymbol{\mu}\boldsymbol{\varphi}-\mathrm{i}\hat{P}(\boldsymbol{\varphi})}}. \tag{4.39}$$

In this form, even though the integral is infinite dimensional in nature, its structure clearly suggest a saddle point calculation. Discarding all subleading terms in $\chi^*[P]$ we get the following result for the leading order in N,

$$\boldsymbol{\varphi}[\hat{\varrho}] = \lim -\int\mathrm{d}\boldsymbol{\varphi}P_s(\boldsymbol{\varphi})\frac{\delta\chi^*}{\delta P(\boldsymbol{\varphi})}[P_s]+\chi^*[P_s]+\log\int\mathrm{d}\boldsymbol{\varphi}\mathrm{e}^{\frac{\mathrm{i}}{2}\boldsymbol{\varphi}\cdot\boldsymbol{\mu}\boldsymbol{\varphi}+\frac{\delta\chi^*}{\delta P(\boldsymbol{\varphi})}[P_s]}. \tag{4.40}$$

Where $P_s(\boldsymbol{\varphi})$ solves the saddle point equation,

$$P_s(\boldsymbol{\varphi}) = \frac{\mathrm{e}^{\frac{\mathrm{i}}{2}\boldsymbol{\varphi}\cdot\boldsymbol{\mu}\boldsymbol{\varphi}+\frac{\delta\chi^*}{\delta P(\boldsymbol{\varphi})}[P_s]}}{\int \mathrm{d}\boldsymbol{\varphi}'\,\mathrm{e}^{\frac{\mathrm{i}}{2}\boldsymbol{\varphi}'\cdot\boldsymbol{\mu}\boldsymbol{\varphi}'+\frac{\delta\chi^*}{\delta P(\boldsymbol{\varphi}')}[P_s]}}. \tag{4.41}$$

This extremely difficult equation is a generalization of the one presented in [59], as it is a general expression and does not assume sparsity or anything besides invariance under node permutation and that the leading order of χ^* is $\mathcal{O}(1)$.

4.2.3 Taking the imaginary replica limit for the replica symmetric ansatz

In order to continue with the calculation it is necessary to introduce a replica symmetric ansatz. We should point out that so far replica symmetry has always been assumed without problems when calculating spectral densities. It has been found that for the calculation of certain higher order observables like eigenvalue correlations one does need to include an ansatz with replica symmetry breaking, [60]. Therefore we introduce the following replica symmetric ansatz,

$$\begin{aligned} P_s(\boldsymbol{\varphi}) =& \mathcal{C}\int \mathrm{d}\mathbf{X}\, W(\mathbf{X})\frac{1}{Z(\mathbf{X})}\mathrm{e}^{-\frac{\mathrm{i}}{2}\boldsymbol{\varphi}\cdot\mathbf{X}\boldsymbol{\varphi}}, \\ Z(\mathbf{X}) =& \int \mathrm{d}\boldsymbol{\varphi}\,\mathrm{e}^{-\frac{\mathrm{i}}{2}\boldsymbol{\varphi}\cdot\mathbf{X}\boldsymbol{\varphi}} = \prod_{\mu}\left(\frac{2\pi}{\mathrm{i}x(\mu)}\right)^{\frac{n_\mu}{2}}\left(\frac{2\pi}{\mathrm{i}\overline{x(\mu)}}\right)^{\frac{m_\mu}{2}}. \end{aligned} \tag{4.42}$$

Where $\mathbf{X}$ is a diagonal matrix,

$$\mathbf{X} = \begin{pmatrix} x(\mu_1)\mathbf{I}_{n_{\mu_1}} & \mathbf{0} & & \dots & & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & & & \\ & \mathbf{0} & x(\mu_M)\mathbf{I}_{n_{\mu_M}} & \mathbf{0} & & \vdots \\ \vdots & & \mathbf{0} & \overline{x(\mu_1)}\mathbf{I}_{m_{\mu_1}} & \mathbf{0} & \\ & & & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & & \dots & & \mathbf{0} & \overline{x(\mu_M)}\mathbf{I}_{m_{\mu_M}} \end{pmatrix}. \tag{4.43}$$

While in general this will only induce an even harder equation for the distribution $W(\mathbf{X})$, we just point out that when evaluating (4.40) all the terms will be functions of $\mathbf{X}$ as all the moments of a Gaussian are determined by the second moments. The resulting expressions

might be quite complicated but in the end one expects to get objects of the following form,

$$\prod_{\mu} f(\mu_\varepsilon)^{n_\mu}\overline{f(\mu_\varepsilon)}^{m_\mu}. \tag{4.44}$$

Whenever a combination like this appears we want to take the imaginary replica limit in the following way,

$$\lim_{\Delta\to 0}\lim_{\substack{n_\mu\to \mathrm{i}\frac{\Delta}{\pi}\hat{\varrho}'(\mu)\\ m_\mu\to -n_\mu}} \exp\left(\sum_\mu n_\mu \log f(\mu) + m_\mu \log\overline{f(\mu)}\right) = \exp\left(\frac{2}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log f(\mu_\varepsilon)\right). \tag{4.45}$$

Note that we keep the notation $\mu_\varepsilon = \mu + \mathrm{i}\varepsilon$ to make explicit the fact that f is a complex function that depends on μ_ε. This is crucial as this is what allows us to define the correct branch of the log in (4.45). As is usual now, we want the derivative of the log to be

$$\frac{\mathrm{d}}{\mathrm{d}\mu}\log f(\mu_\varepsilon) = \frac{f'(\mu_\varepsilon)}{f(\mu_\varepsilon)}. \tag{4.46}$$

This might seem trivial but it is highly nontrivial since, given that n_μ and m_μ are integers, it is always possible to add an extra integer phase without modifying the identity,

$$\prod_{\mu} f(\mu_\varepsilon)^{n_\mu}\overline{f(\mu_\varepsilon)}^{m_\mu} = \exp\left(\sum_\mu n_\mu \log f(\mu_\varepsilon) + m_\mu \log\overline{f(\mu_\varepsilon)} + 2\pi\mathrm{i}n_\mu r(\mu) + 2\pi\mathrm{i}m_\mu t(\mu)\right). \tag{4.47}$$

Where $r(\mu), t(\mu) \in \mathbb{Z}$. If we then take the replica limit we would get,

$$\lim\prod_{\mu} f(\mu_\varepsilon)^{n_\mu}\overline{f(\mu_\varepsilon)}^{m_\mu} = \exp\left(\frac{2}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log f(\mu_\varepsilon) + 4\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}[r(\mu)-t(\mu)]\right) \tag{4.48}$$

If we now take the functional derivative w.r.t $\hat{\varrho}(\mu)$ we get

$$\mathrm{e}^{\frac{2}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log f(\mu_\varepsilon)+4\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}[r(\mu)-t(\mu)]}\left[\frac{2}{\pi}\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\frac{f'(\mu_\varepsilon)}{f(\mu_\varepsilon)} + 4\frac{\mathrm{d}}{\mathrm{d}\mu}[r(\mu)-t(\mu)]\right] \tag{4.49}$$

And the crucial observation is that given that both $r(\mu), t(\mu)$ are integers, its derivative in (4.49) should be 0 or a δ function. But it cannot be a δ function because at finite $\varepsilon > 0$ the functional derivative should generate imaginary parts of analytic functions, as pointed

out before in (4.10). Therefore it makes sense to keep only the first term for consistency.

Our conjecture is then that the condition (4.46) defines the replica symmetric solution in a unique way without having to worry about the issues with analytic continuation. Although unlikely, it could be that f alone is not the correct function that gives the correct analytic derivative. This sort of issue has never been observed and is the main reason why it is rarely discussed. If it were the case that (4.46) gave the wrong answer then our conjecture is that the following recipe is the correct one for the replica limit.

In summary, our conjecture is,

- When encountering an expression of the form $\lim \prod_\mu f(\mu_\varepsilon)^{n_\mu} \overline{f(\mu_\varepsilon)}^{m_\mu}$, the correct replica limit is

$$\exp\left(\frac{2}{\pi}\int \mathrm{d}\mu \hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log_a f(\mu_\varepsilon)\right). \tag{4.50}$$

  Where $\log_a f(\mu_\varepsilon)$ means taking the branch such that $\frac{\mathrm{d}}{\mathrm{d}\mu}\log_a f(\mu_\varepsilon)$ is an analytic function of μ_ε.

With this conjecture, (4.40) can be calculated and the functional derivative can be calculated. The replica limit always gives the following simple expression, from which to take the functional derivatives:

$$\frac{\delta}{\delta\hat{\varrho}(\mu)}\left[\mathrm{e}^{\int \mathrm{d}\mu' \hat{\varrho}(\mu')a(\mu')}\right] = \mathrm{e}^{\int \mathrm{d}\mu' \hat{\varrho}(\mu')a(\mu')}a(\mu). \tag{4.51}$$

Note that we can also analyze the traditional replica limit in this light and conjecture it should be taken in the same way,

$$\lim_{n\to 0}\frac{[f(\mu_\varepsilon)]^n - 1}{n} = \log_a f(\mu_\varepsilon) \text{ such that } \frac{\mathrm{d}}{\mathrm{d}\mu}\log_a f(\mu_\varepsilon) \text{ is analytic} \tag{4.52}$$

So far all authors using the replica method for spectral densities of finitely connected graphs have worked with this choice, either consciously or unconsciously.

4.3 Discussion

We have shown a way to solve analytically a functional ensemble that allows to calculate the following two different objects in a unified way.

1. The generating function for a particular model when evaluating the generating function for a specific $\hat{\varrho}(\mu)$, for example $\hat{\varrho}(\mu) = \alpha\mu^3$

2. The average spectral density for that model when substituting the specific $\hat{\varrho}(\mu)$ in the functional derivative $\delta\phi[\hat{\rho}]/\delta\hat{\varrho}(\mu)$

In the subsequent chapters we will work out the theory for the particular case where the unbiased distribution is the configuration model (CM) of unweighted simple graphs.

$$\begin{aligned} p_0(\mathbf{A}) =& \frac{1}{\mathcal{N}_\mathbf{k}} \prod_{i=1}^{N} \delta_{k_i, \sum_j A_{ij}} \\ \mathcal{N}_\mathbf{k} =& \sum_{\mathbf{A}} \prod_{i=1}^{N} \delta_{k_i, \sum_j A_{ij}} \end{aligned} \tag{4.53}$$

Along the way we have also clarified certain doubts on what is the correct way to take the imaginary replica limit and what it means. Although it had not really been an issue in literature so far, in Chapter 6 we will show certain expressions where it *is* necessary to choose the correct branch *by hand* or otherwise get incorrect results, see the discussion below (6.49).

The biggest question for this general approach is how much one can deviate from the original untilted ensemble $p_0(\boldsymbol{H})$. Looking at equation (4.41) we can see that it heavily depends on $\chi^*[P]$ which only depends on $p_0(\boldsymbol{H})$ and not on $\hat{\varrho}$, the dependence on $\hat{\varrho}$ is recovered when taking the replica limit for the saddle point equation itself. But it could be that the initial use $\chi^*[P]$ has left out valuable contributions. For the case of (4.53) and $\hat{\varrho}(\mu) = \alpha\mu^3$ this will not be so.

Chapter 5

Imaginary replica analysis of loopy regular random graphs

In this chapter we introduce a random graph ensemble of random regular graphs with a spectral constraint. In order to work with it we use the imaginary replica trick. The regularity of the degree distribution will allow for an exact solution of the theory up to $\mathcal{O}(1/N)$.

5.1 Definitions

We study ensembles of simple nondirected N-node regular graphs with degree q. Each graph is defined by its symmetric $N \times N$ adjacency matrix $\mathbf{A}$, with entries $A_{ij} = 1$ if nodes i and j are connected, and $A_{ij} = 0$ otherwise ($A_{ii} = 0$ for all i). The formula for the total number of closed paths of a given length in a graph $\mathbf{A}$ is, apart from a simple overcounting factor,

$$\mathrm{Tr}(\mathbf{A}^{\ell}) \quad = \quad \#\text{ of closed paths of length } \ell. \tag{5.1}$$

It follows from the relation $\mathrm{Tr}(\mathbf{A}^{\ell}) = N \int \mathrm{d}\mu\, \varrho(\mu|\mathbf{A})\mu^{\ell}$. that controlling the numbers of closed paths of all lengths ℓ in random graphs is equivalent to controlling the moments of the spectral density $\varrho(\mu|\mathbf{A})$ of $\mathbf{A}$,

$$\varrho(\mu|\mathbf{A}) = \frac{1}{N}\sum_{i=1}^{N} \delta(\mu - \mu_i(\mathbf{A})), \tag{5.2}$$

where $\mu_i(\mathbf{A})$ is the i-th eigenvalue of $\mathbf{A}$.

In exponential spectrally constrained ensembles [47], the graph probabilities $p(\mathbf{A})$ on the set G of simple nondirected N-node graphs are determined by maximising the Shannon entropy $S[p] = -\sum_{\mathbf{A}\in G} p(\mathbf{A}) \log p(\mathbf{A})$, subject to prescribed values of all degrees and a prescribed expectation value of the spectral density. For q-regular random graphs this gives

$$p(\mathbf{A}) = \frac{\mathrm{e}^{N\int \mathrm{d}\mu\ \hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})}}{Z[\hat{\varrho}]} \prod_{i=1}^{N} \delta_{q,\sum_j A_{ij}}, \tag{5.3}$$

$$Z[\hat{\varrho}] = \sum_{\mathbf{A}\in G} \mathrm{e}^{N\int \mathrm{d}\mu\ \hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})} \prod_{i=1}^{N} \delta_{q,\sum_j A_{ij}}. \tag{5.4}$$

Here $\hat{\varrho}(\mu)$ is a functional Lagrange multiplier. By construction, (5.3) defines the most unbiased ensemble of q-regular nondirected graphs with a prescribed adjacency matrix spectrum. We write averages over (5.3) as $\langle f(\mathbf{A})\rangle = \sum_{\mathbf{A}\in G} p(\mathbf{A}) f(\mathbf{A})$. The expected eigenvalue density $\varrho(\mu) = \langle \varrho(\mu|\mathbf{A})\rangle$ can be obtained from a generating function $\phi[\hat{\varrho}]$:

$$\varrho(\mu) = \frac{\delta\phi[\hat{\varrho}]}{\delta\hat{\varrho}(\mu)}, \qquad \phi[\hat{\varrho}] = N^{-1}\log Z[\hat{\varrho}]. \tag{5.5}$$

Our main interest is in finding an analytical expression for the expected density $\varrho(\mu)$ in terms of the functional Lagrange parameter $\hat{\varrho}(\mu)$.

For the simple choice $\hat{\varrho}(\mu) = \alpha\mu^3$ we recover from (5.3) the model of [27], in which the number of loops of length three (i.e. of triangles) is constrained:

$$p(\mathbf{A}) = \frac{\mathrm{e}^{\alpha \mathrm{Tr}(\mathbf{A}^3)}}{Z(\alpha)} \prod_{i=1}^{N} \delta_{q,\sum_j A_{ij}}, \qquad Z(\alpha) = \sum_{\mathbf{A}\in G} \mathrm{e}^{\alpha \mathrm{Tr}(\mathbf{A}^3)} \prod_{i=1}^{N} \delta_{q,\sum_j A_{ij}}. \tag{5.6}$$

The number of triangles $n_3(\mathbf{A})$ differs from the trace only by an overcounting factor, viz. $n_3(\mathbf{A}) = \frac{1}{6}\mathrm{Tr}(\mathbf{A}^3)$. The generating function associated with (5.6) is $\phi(\alpha) = N^{-1}\log Z(\alpha)$. From $m(\alpha) = \partial\phi(\alpha)/\partial\alpha = N^{-1}\langle \mathrm{Tr}(\mathbf{A}^3)\rangle$ follows the average clustering coefficient $\langle C(\mathbf{A})\rangle = m(\alpha)/q(q-1)$. For $q = 2$ one can compute $\phi(\alpha)$ using standard combinatorics [27], for both $\alpha = \mathcal{O}(1)$ and $\alpha = \mathcal{O}(\log N)$. For arbitrary q, the expected loop density $m(\alpha)$ always vanishes for $N \to \infty$ if $\alpha = \mathcal{O}(1)$. As in [27], one could rescale α with a factor $\log N$. While this could give an asymptotic theory with finite loop densities, we will find that it would not be a useful network model for applications.

5.2 Evaluation of the generating function

5.2.1 Imaginary replica approach

We note that all q-regular graphs of size N have identical probabilities in an Erdös-Renyi (ER) ensemble with average degree q,

$$p_{\rm ER}(\mathbf{A}) = \Big(\frac{q}{N}\Big)^{\frac{Nq}{2}} \Big(\frac{N-q}{N}\Big)^{\frac{N}{2}(N-1-q)} \quad \text{for all } \mathbf{A}\in G \ \text{ with } \ \prod_{i=1}^{N} \delta_{q,\sum_j A_{ij}} = 1. \tag{5.7}$$

Hence we can rewrite (5.5) as an average over this ER ensemble:

$$\phi[\hat{\varrho}] \ = \ \frac{1}{N} \log \Big\langle {\rm e}^{N\int {\rm d}\mu \ \hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})} \prod_{i=1}^{N} \delta_{q,\sum_j A_{ij}} \Big\rangle_{\rm ER} + \text{constant}. \tag{5.8}$$

Upon using the Edwards-Jones formula [36] for $\varrho(\mu|\mathbf{A})$, writing the integral over eigenvalues as $\int {\rm d}\mu \ldots = \lim_{\Delta\to 0} \Delta \sum_\mu \ldots$, and after some modest manipulations, the key quantity in this expression can be written as follows, with infinitely many imaginary replicas (two for each eigenvalueμ, n_μ and m_μ):

$${\rm e}^{N\int {\rm d}\mu \ \hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})} \ = \ \lim_{\Delta\to 0} \lim_{n_\mu \to {\rm i}\frac{\Delta}{\pi}\hat{\varrho}'(\mu)} \lim_{m_\mu\to -n_\mu} \prod_\mu Z(\mu_\varepsilon|\mathbf{A})^{n_\mu} \overline{Z(\mu_\varepsilon|\mathbf{A})}^{m_\mu}. \tag{5.9}$$

Here $\mu_\varepsilon = \mu + {\rm i}\varepsilon$ and

$$Z(\mu_\varepsilon|\mathbf{A}) \ = \ \int \prod_{i=1}^{N} {\rm d}\phi^i \exp\Big[-\frac{\rm i}{2} \sum_{ij} \phi^i (A_{ij} - \mu_\varepsilon \delta_{ij}) \phi^j\Big]. \tag{5.10}$$

One initially takes $n_\mu, m_\mu \in \mathbb{N}$, in order to perform the calculation, followed by analytic continuation to the relevant imaginary values. In its above form, (5.9) appeared first in [47, 52], but similar formulae involving limits of replica dimensions to non-zero values have been introduced in different contexts, in particular when counting the number of eigenvalues in certain intervals for random matrix ensembles, see e.g. [54–58, 61, 62]. We can combine the integrals in (5.9) as follows:

$$\begin{aligned}\prod_\mu Z(\mu_\varepsilon|\mathbf{A})^{n_\mu} \overline{Z(\mu_\varepsilon|\mathbf{A})}^{m_\mu} = \int {\rm d}[\mathbf{\Phi},\mathbf{\Psi}] \exp\Big(-\frac{\rm i}{2} \sum_{\mu,\alpha_\mu} \sum_{ij} \phi^i_{\mu,\alpha_\mu} \phi^j_{\mu,\alpha_\mu} (A_{ij} - \mu_\varepsilon \delta_{ij})\Big) \\ \times \exp\Big(\frac{\rm i}{2} \sum_{\mu,\beta_\mu} \sum_{ij} \psi^i_{\mu,\beta_\mu} \psi^j_{\mu,\beta_\mu} (A_{ij} - \overline{\mu_\varepsilon} \delta_{ij})\Big),\end{aligned} \tag{5.11}$$

where $d[\mathbf{\Phi},\mathbf{\Psi}]=\prod_{i=1}^{N}\prod_{\mu}\left[\left(\prod_{\alpha_\mu=1}^{n_\mu}\mathrm{d}\boldsymbol{\phi}^i_{\mu,\alpha_\mu}\right)\left(\prod_{\beta_\mu=1}^{m_\mu}\mathrm{d}\psi^i_{\mu,\beta_\mu}\right)\right]$. To simplify our notation we introduce the vector $\boldsymbol{\varphi}\in\mathbb{R}^d$, where $d=\sum_\mu n_\mu+\sum_\mu m_\mu$, with entries

$$\boldsymbol{\varphi}=\begin{pmatrix}\{\phi_{\mu,\alpha_\mu}\}\\ \{\psi_{\mu,\beta_\mu}\}\end{pmatrix}, \tag{5.12}$$

the dot product

$$\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'=\sum_{\mu,\alpha_\mu}\phi_{\mu,\alpha_\mu}\phi'_{\mu,\alpha_\mu}-\sum_{\mu,\beta_\mu}\psi_{\mu,\beta_\mu}\psi'_{\mu,\beta_\mu}, \tag{5.13}$$

and the following $d\times d$ diagonal matrix $\boldsymbol{\mu}$, in which $\mu\in\{\mu_1,\ldots,\mu_M\}$ and $\mathbf{I}_n$ denotes the n dimensional identity matrix:

$$\boldsymbol{\mu}=\begin{pmatrix}
\mu_{1\varepsilon}\mathbf{I}_{n_{\mu_1}} & \mathbf{0} & & \ldots & & \mathbf{0}\\
\mathbf{0} & \ddots & \mathbf{0} & & & \\
 & \mathbf{0} & \mu_{M\varepsilon}\mathbf{I}_{n_{\mu_M}} & \mathbf{0} & & \vdots\\
\vdots & & \mathbf{0} & \overline{\mu_{1\varepsilon}}\mathbf{I}_{m_{\mu_1}} & \mathbf{0} & \\
 & & & \mathbf{0} & \ddots & \mathbf{0}\\
\mathbf{0} & & \ldots & & \mathbf{0} & \overline{\mu_{M\varepsilon}}\mathbf{I}_{m_{\mu_M}}
\end{pmatrix}. \tag{5.14}$$

Here M denotes the number of μ-values in the discretized eigenvalue integral, so $M\to\infty$ when we take the limit $\Delta\to 0$. We finally introduce the two shorthands

$$\nu(\boldsymbol{\varphi})=\mathrm{e}^{\frac{1}{2}\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\mu}\boldsymbol{\varphi}},\qquad \lim=\lim_{\Delta\to 0}\lim_{n_\mu\to\mathrm{i}\frac{\Delta}{\pi}\hat{\varrho}'(\mu)}\lim_{m_\mu\to-n_\mu}. \tag{5.15}$$

The above definitions, together with the integral form of the Kronecker delta, $\delta_{nm}=(2\pi)^{-1}\int_{-\pi}^{\pi}\mathrm{d}\omega\ \exp[\mathrm{i}\omega(n-m)]$, enable us to compute the generating function (5.8) following [47], which in turn is reminiscent of previous spectral calculations for sparse random graphs [39, 63]. Upon dropping the irrelevant constant in (5.8) we get

$$\begin{aligned}\phi[\hat{\varrho}] &= \lim\frac{1}{N}\log\int\prod_{i=1}^{N}\left[\mathrm{d}\boldsymbol{\varphi}^i\ \nu(\boldsymbol{\varphi}^i)\frac{\mathrm{d}\omega_i}{2\pi}\,\mathrm{e}^{\mathrm{i}\omega_i q}\right]\left\langle\mathrm{e}^{-\mathrm{i}\sum_{i<j}A_{ij}[\boldsymbol{\varphi}^i\cdot\boldsymbol{\varphi}^j+\omega_i+\omega_j]}\right\rangle_{\mathrm{ER}}\\ &= \lim\frac{1}{N}\log\int\prod_{i=1}^{N}\left[\mathrm{d}\boldsymbol{\varphi}^i\ \nu(\boldsymbol{\varphi}^i)\frac{\mathrm{d}\omega_i}{2\pi}\,\mathrm{e}^{\mathrm{i}\omega_i q}\right]\mathrm{e}^{\sum_{i<j}\log\left[1+\frac{q}{N}\left(\mathrm{e}^{-\mathrm{i}[\boldsymbol{\varphi}^i\cdot\boldsymbol{\varphi}^j+\omega_i+\omega_j]}-1\right)\right]},\end{aligned} \tag{5.16}$$

Since we intend to compute finite size spectrum fluctuations, in expanding the logarithm for large N we keep both the $\mathcal{O}(N)$ and $\mathcal{O}(1)$ terms in the exponent:

$$\begin{aligned}\phi[\hat{\varrho}] &= -\frac{q}{2}(1+\frac{q-2}{2N}) + \lim \frac{1}{N}\log\int\prod_{i=1}^{N}\Big[\mathrm{d}\boldsymbol{\varphi}^i\nu(\boldsymbol{\varphi}^i)\frac{\mathrm{d}\omega_i}{2\pi}\,\mathrm{e}^{\mathrm{i}\omega_i q-\frac{q}{2N}\mathrm{e}^{-\mathrm{i}[\boldsymbol{\varphi}^i\cdot\boldsymbol{\varphi}^i+2\omega_i]}}\Big] \qquad (5.17)\\ &\quad\times\exp\Big[\frac{q}{2N}(1+\frac{q}{N})\sum_{ij}\mathrm{e}^{-\mathrm{i}[\boldsymbol{\varphi}^i\cdot\boldsymbol{\varphi}^j+\omega_i+\omega_j]}-\frac{q^2}{4N^2}\sum_{ij}\mathrm{e}^{-2\mathrm{i}[\boldsymbol{\varphi}^i\cdot\boldsymbol{\varphi}^j+\omega_i+\omega_j]}\Big]+\mathcal{O}(\frac{1}{N^2}).\end{aligned}$$

We now introduce the order parameter

$$P(\boldsymbol{\varphi},\omega)=\frac{1}{N}\sum_{i=1}^{N}\delta(\boldsymbol{\varphi}-\boldsymbol{\varphi}^i)\delta(\omega-\omega_i) \qquad (5.18)$$

We enforce it by inserting the following functional integral, obtained by writing delta functions for each $(\boldsymbol{\varphi},\omega)$ in integral representation, and with the usual path integral measure $\mathcal{D}P=\prod_{\boldsymbol{\phi}}\prod_{\omega}[\mathrm{d}P(\boldsymbol{\phi},\omega)\sqrt{N\Delta_\varphi\Delta_\omega/2\pi})]$ (where $\Delta_\varphi,\Delta_\omega\to 0$):

$$1=\int\mathcal{D}P\mathcal{D}\hat{P}\;\mathrm{e}^{N\mathrm{i}\int\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\omega P(\boldsymbol{\varphi},\omega)\hat{P}(\boldsymbol{\varphi},\omega)-\mathrm{i}\int\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\omega\hat{P}(\boldsymbol{\varphi},\omega)\sum_i\delta(\boldsymbol{\varphi}-\boldsymbol{\varphi}^i)\delta(\omega-\omega_i)}. \qquad (5.19)$$

The result is:

$$\phi[\hat{\varrho}] = \lim\frac{1}{N}\log\int\mathcal{D}P\mathcal{D}\hat{P}\;\mathrm{e}^{NS[P,\hat{P}]}+\mathcal{O}(\frac{1}{N^2}), \qquad (5.20)$$

$$\begin{aligned}S[P,\hat{P}] &= -\frac{q}{2}(1+\frac{q-2}{2N})+\mathrm{i}\int\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\omega\;P(\boldsymbol{\varphi},\omega)\hat{P}(\boldsymbol{\varphi},\omega)\\ &\quad+\frac{q}{2}(1+\frac{q}{N})\int\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\omega\mathrm{d}\omega'\;P(\boldsymbol{\varphi},\omega)P(\boldsymbol{\varphi}',\omega')\mathrm{e}^{-\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'-\mathrm{i}\omega-\mathrm{i}\omega'}\\ &\quad-\frac{q^2}{4N}\int\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\omega\mathrm{d}\omega'\;P(\boldsymbol{\varphi},\omega)P(\boldsymbol{\varphi}',\omega')\mathrm{e}^{-2\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'-2\mathrm{i}\omega-2\mathrm{i}\omega'} \qquad (5.21)\\ &\quad-\frac{q}{2N}\int\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\omega\;P(\boldsymbol{\varphi},\omega)\mathrm{e}^{-\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}-2\mathrm{i}\omega}+\log\int\frac{\mathrm{d}\omega\mathrm{d}\boldsymbol{\varphi}}{2\pi}\nu(\boldsymbol{\varphi})\mathrm{e}^{\mathrm{i}\omega q-\mathrm{i}\hat{P}(\boldsymbol{\varphi},\omega)}.\end{aligned}$$

It was shown in [63] how this type of integral can be reduced to an integral over a single functional variable. In C.1 we work out the details, leading to

$$\phi[\hat{\varrho}] = \lim\Big\{\log\int\mathrm{d}\boldsymbol{\varphi}\;\nu(\boldsymbol{\varphi})\Big[\int\mathrm{d}\boldsymbol{\varphi}'\;U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W_0(\boldsymbol{\varphi}')\Big]^q+\frac{1}{2N}\sum_{\ell=3}^{\infty}\frac{\mathrm{Tr}(\mathbf{T}^\ell)}{\ell}\Big\}, \qquad (5.22)$$

in which $U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}') = \mathrm{e}^{-\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'}$, and the function $W_0(\boldsymbol{\varphi})$ is to be solved from

$$W_0(\boldsymbol{\varphi}) = \frac{\nu(\boldsymbol{\varphi})}{Z_q}\Big[\int \mathrm{d}\boldsymbol{\varphi}'\; U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W_0(\boldsymbol{\varphi}')\Big]^{q-1}, \tag{5.23}$$

$$Z_q = \int \mathrm{d}\boldsymbol{\varphi}\; \nu(\boldsymbol{\varphi})\Big[\int \mathrm{d}\boldsymbol{\varphi}'\; U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W_0(\boldsymbol{\varphi}')\Big]^{q}, \tag{5.24}$$

and

$$T(\boldsymbol{\varphi},\boldsymbol{\varphi}') = (q-1)r[W_0(\boldsymbol{\varphi})]U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}') - qW_0(\boldsymbol{\varphi})\int \mathrm{d}\boldsymbol{\psi}\; U_1(\boldsymbol{\varphi}',\boldsymbol{\psi})W_0(\boldsymbol{\psi}), \tag{5.25}$$

$$r[W_0(\boldsymbol{\varphi})] = \frac{\nu(\boldsymbol{\varphi})}{Z_q}\Big[\int \mathrm{d}\boldsymbol{\varphi}'\; U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W_0(\boldsymbol{\varphi}')\Big]^{q-2}. \tag{5.26}$$

Finally, following [64] we may use the following identity[1]:

$$\mathrm{Tr}(\mathbf{T}^\ell) = (q-1)^\ell[\mathrm{Tr}(\mathbf{M}^\ell)-1] + (-1)^\ell \tag{5.27}$$

with $M(\boldsymbol{\varphi},\boldsymbol{\varphi}') = r[W_0(\boldsymbol{\varphi})]U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')$, to simplify (5.22) modulo an additive constant to

$$\phi[\hat{\varrho}] = \lim\Big\{\log Z_q + \sum_{\ell=3}^{\infty}\frac{(q-1)^\ell}{2N\ell}\mathrm{Tr}(\mathbf{M}^\ell)\Big\}. \tag{5.28}$$

Expression (5.28) was originally presented in [63], and we have indeed chosen our notation at the start deliberately to emphasize and exploit the similarity. However, although identical in structure, the present formula (5.28) differs from the one in [63] in the underlying definitions of the fields $\boldsymbol{\varphi}$, the dot product $\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'$ and the function $\nu(\boldsymbol{\varphi})$, which here all involve the full eigenvalue spectrum and describe our present controlled non-uniform measures over the space of graphs.

5.2.2 Replica symmetric solution

In order to continue, we assume that the order parameter $W_0(\boldsymbol{\varphi})$ is replica symmetric (RS), i.e. invariant under all permutations of all replicas (noting that in the present problem we have a separate replica index for each eigenvalue μ, and that $W_0(\boldsymbol{\varphi})$ is not a normalized distribution). The RS assumption is necessary to carry on with the calculation formally. For the present type of spectral calculations, RS has in the past always shown accurate results, [39, 54, 55, 57, 58, 61–63]. Many of these RS analyses reproduced

[1] The proof of this interesting identity follows directly from the operator properties $\mathbf{MB} = \mathbf{BM} = \mathbf{B}^2 = \mathbf{I}$, where $M(\boldsymbol{\varphi},\boldsymbol{\varphi}') = r[W_0(\boldsymbol{\varphi})]U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')$ and $B(\boldsymbol{\varphi},\boldsymbol{\varphi}') = W_0(\boldsymbol{\varphi})\int \mathrm{d}\boldsymbol{\psi}\; U_1(\boldsymbol{\varphi}',\boldsymbol{\psi})W_0(\boldsymbol{\psi})$.

rigorous expressions for both the asymptotic spectral density, [46, 65], and the finite size fluctuations, [66]. The need to break replica symmetry has only been observed so far when calculating higher order objects like average correlations of eigenvalues in [62], therefore we do not see a need to consider replica symmetry breaking at this stage to get exact results. To represent $W_0(\varphi)$ in a RS way we choose a superposition of zero mean complex Gaussian distributions, following [37, 39] where it was shown that this family of distributions is appropriate and natural for a quadratic interaction among the variables $\boldsymbol{\varphi}$, such as in (5.10). The exactness of our final results will validate this specific choice *a posteriori* over other possible RS approaches. Therefore we use the following ansatz:

$$W_0(\boldsymbol{\varphi}) = \mathcal{C}\int \mathrm{d}\mathbf{X}\; W(\mathbf{X})\frac{1}{Z(\mathbf{X})}\mathrm{e}^{-\frac{1}{2}\mathrm{i}\boldsymbol{\varphi}\cdot\mathbf{X}\boldsymbol{\varphi}}, \tag{5.29}$$

$$Z(\mathbf{X}) = \int \mathrm{d}\boldsymbol{\varphi}\; \mathrm{e}^{-\frac{1}{2}\mathrm{i}\boldsymbol{\varphi}\cdot\mathbf{X}\boldsymbol{\varphi}} = \prod_{\mu}\Big(\frac{2\pi}{\mathrm{i}x(\mu)}\Big)^{\frac{n_\mu}{2}}\Big(\frac{\overline{2\pi}}{\overline{\mathrm{i}x(\mu)}}\Big)^{\frac{m_\mu}{2}}, \tag{5.30}$$

where $\mathbf{X}\in\mathbb{R}^{d\times d}$ is a diagonal matrix with the following structure:

$$\mathbf{X} = \begin{pmatrix} x(\mu_1)\mathbf{I}_{n_{\mu_1}} & \mathbf{0} & & \dots & & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & & & \\ & \mathbf{0} & x(\mu_M)\mathbf{I}_{n_{\mu_M}} & \mathbf{0} & & \vdots \\ \vdots & & \mathbf{0} & \overline{x(\mu_1)}\mathbf{I}_{m_{\mu_1}} & \mathbf{0} & \\ & & & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & & \dots & & \mathbf{0} & \overline{x(\mu_M)}\mathbf{I}_{m_{\mu_M}} \end{pmatrix}. \tag{5.31}$$

Expression (5.30) is indeed invariant under all permutations of replica indices with any fixed value of μ, that is $\{\phi_{\mu,1},\dots,\phi_{\mu,n_\mu}\}$ and $\{\psi_{\mu,1},\dots,\psi_{\mu,m_\mu}\}$. The new RS order parameter is the distribution $W(\mathbf{X})$, where each $\mathbf{X}$ is specified by M complex numbers $x(\mu)$. Hence the integration in (5.13) is over the real and imaginary part of each $x(\mu)$, so $\mathrm{d}\mathbf{X} = \prod_\mu \mathrm{d}\,\mathrm{Re}[x(\mu)]\mathrm{d}\,\mathrm{Im}[x(\mu)]$. For (5.29) to be well defined, we must restrict all $x(\mu)$ to have $\mathrm{Im}\;x(\mu)<0$. We may assume that $\int\mathrm{d}\mathbf{X}\;W(\mathbf{X})=1$, since possible non-normalization of W_0 is reflected in the inclusion in (5.30) of a constant $\mathcal{C}$.

We note that for the present definition (5.13) of the dot product, the following identity is still valid.

$$\int \mathrm{d}\boldsymbol{\varphi}'\; \mathrm{e}^{-\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'-\frac{1}{2}\mathrm{i}\boldsymbol{\varphi}'\cdot\mathbf{X}\boldsymbol{\varphi}'} = Z(\mathbf{X})\mathrm{e}^{\frac{1}{2}\mathrm{i}\boldsymbol{\varphi}\cdot\mathbf{X}^{-1}\boldsymbol{\varphi}}. \tag{5.32}$$

Insertion of our RS ansatz (5.30) into the full order parameter equation (5.23) shows, using the above identity, that the RS ansatz indeed gives a solution of (5.23), provided the RS order parameter satisfies

$$W(\mathbf{X}) = \frac{Z(\mathbf{X})}{\mathcal{Z}_{q-1}} \int \Big(\prod_{k=1}^{q-1} \mathrm{d}\mathbf{X}_k W(\mathbf{X}_k) \Big)\ \delta\Big(\mathbf{X} + \boldsymbol{\mu} + \sum_{k=1}^{q-1} \mathbf{X}_k^{-1}\Big), \tag{5.33}$$

$$\mathcal{Z}_q = \int \mathrm{d}\mathbf{X} \Big(\prod_{k=1}^{q} \mathrm{d}\mathbf{X}_k W(\mathbf{X}_k) \Big)\ Z(\mathbf{X})\ \delta\Big(\mathbf{X} + \boldsymbol{\mu} + \sum_{k=1}^{q} \mathbf{X}_k^{-1}\Big), \tag{5.34}$$

$$\mathcal{C}^2 = \mathcal{Z}_{q-1}/\mathcal{Z}_q. \tag{5.35}$$

We similarly derive $Z_q = \mathcal{C}^q \mathcal{Z}_q$, giving in combination with (5.35):

$$\log Z_q = \frac{1}{2} q \log \mathcal{Z}_{q-1} - \frac{1}{2}(q-2) \log \mathcal{Z}_q. \tag{5.36}$$

The above RS order parameter equations (5.33,5.34) have one specific simple solution, namely the delta distribution $W(\mathbf{X}) = \delta(\mathbf{X} - \mathbf{X}^\star)$, in which the entries of $\mathbf{X}^\star$ satisfy

$$x^\star(\mu) = -\mu_\varepsilon - \frac{q-1}{x^\star(\mu)}. \tag{5.37}$$

Of the two possible solutions of this equation we must choose the one with $\mathrm{Im}\ x(\mu) < 0$:

$$x^\star(\mu) = -\tfrac{1}{2}\mu_\varepsilon - \tfrac{1}{2}\mathrm{i}\sqrt{4(q-1) - {\mu_\varepsilon}^2}. \tag{5.38}$$

For this special RS solution we have

$$W_0(\boldsymbol{\varphi}) = \frac{\mathcal{C}}{Z(\mathbf{X}^\star)} \mathrm{e}^{-\frac{1}{2}\mathrm{i}\boldsymbol{\varphi}\cdot\mathbf{X}^\star\boldsymbol{\varphi}}, \tag{5.39}$$

$$\mathcal{Z}_{q-1} = Z(\mathbf{X}^\star), \tag{5.40}$$

$$\mathcal{Z}_q = \int \mathrm{d}\mathbf{X}\ Z(\mathbf{X})\ \delta\Big(\mathbf{X} + \boldsymbol{\mu} + q(\mathbf{X}^\star)^{-1}\Big), \tag{5.41}$$

$$= Z(\mathbf{X}^\star - (\mathbf{X}^\star)^{-1}),$$

and hence also $\mathcal{C}^2 = Z(\mathbf{X}^\star)/Z(\mathbf{X}^\star - (\mathbf{X}^\star)^{-1})$. The kernel $\mathbf{M}$ which appears in the generating function (5.28) will now have the following entries:

$$M(\boldsymbol{\varphi}, \boldsymbol{\varphi}') = \mathcal{Z}_{q-1}^{-1}\ \mathrm{e}^{\frac{1}{2}\mathrm{i}\boldsymbol{\varphi}\cdot[\boldsymbol{\mu} + (q-2)(\mathbf{X}^\star)^{-1}]\boldsymbol{\varphi} - \mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'}. \tag{5.42}$$

In C.2 we work out the traces $\mathrm{Tr}(\mathbf{M}^\ell)$ for the RS solution, and find

$$\mathrm{Tr}(\mathbf{M}^\ell) \;=\; \mathcal{Z}_{q-1}^{-\ell}\prod_\mu\Big[Z(\mu_\varepsilon|\mathbf{A}^\star_{\ell,\mu})^{n_\mu}\overline{Z(\mu_\varepsilon|\mathbf{A}^\star_{\ell,\mu})}^{m_\mu}\Big]. \tag{5.43}$$

Where $Z(\mu_\varepsilon|\mathbf{A}^\star_{\ell,\mu})$ denotes the original complex Gaussian integral defined in (5.10), and $\mathbf{A}^\star_{\ell,\mu}$ is now the $\ell\times\ell$ adjacency matrix of a loop of length ℓ in the presence of a complex field acting on the diagonal, of value $(2-q)/x^\star(\mu)$:

$$\big(\mathbf{A}^\star_{\ell,\mu}\big)_{kk'} \;=\; \delta_{k,k'+1}+\delta_{k,k'-1}+\frac{2-q}{x^\star(\mu)}\delta_{kk'} \qquad (\text{with } k \bmod \ell). \tag{5.44}$$

Substituting (5.36) and (5.44) into expression (5.28) for the generating function, followed by using formulae (5.40,5.42) for the constants $\mathcal{Z}_{q-1}$ and $\mathcal{Z}_q$, then gives

$$\begin{aligned}\phi[\hat\varrho] \;=\; &\lim\Big\{\frac{1}{2}q\log Z(\mathbf{X}^\star)-\frac{1}{2}(q-2)\log Z(\mathbf{X}^\star-(\mathbf{X}^\star)^{-1})\\ &+\sum_{\ell=3}^\infty\frac{(q-1)^\ell}{2N\ell}\frac{1}{Z^\ell(\mathbf{X}^\star)}\prod_\mu\Big[Z(\mu_\varepsilon|\mathbf{A}^\star_{\ell,\mu})^{n_\mu}\overline{Z(\mu_\varepsilon|\mathbf{A}^\star_{\ell,\mu})}^{m_\mu}\Big]\Big\}.\end{aligned} \tag{5.45}$$

5.2.3 Imaginary replica limits

At this stage we can safely take the three limits defined in (5.15). For this we follow the discussion of the previous chapter:

$$\begin{aligned}\lim\prod_\mu f(\mu)^{n_\mu}\overline{f(\mu)}^{m_\mu} \;&=\; \lim_{\Delta\to0}\lim_{n_\mu\to\mathrm{i}\frac{\Delta}{\pi}\hat\varrho'(\mu)}\lim_{m_\mu\to-n_\mu}\mathrm{e}^{\sum_\mu\left[n_\mu\log f(\mu)+m_\mu\log\overline{f(\mu)}\right]}\\ &=\; \mathrm{e}^{\frac{2}{\pi}\int\mathrm{d}\mu\;\hat\varrho(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log f(\mu)}.\end{aligned} \tag{5.46}$$

Where log is defined as in (4.46),

$$\frac{\mathrm{d}}{\mathrm{d}\mu}\log f(\mu_\varepsilon)=\frac{f'(\mu_\varepsilon)}{f(\mu_\varepsilon)} \tag{5.47}$$

In particular, application to $f(\mu)=[2\pi/\mathrm{i}x^\star(\mu)]^{\frac{1}{2}}$ and to $f(\mu)=Z(\mu_\varepsilon|\mathbf{A}^\star_{\ell,\mu})$ gives

$$\lim\log Z(\mathbf{X}^*) \;=\; -\frac{1}{\pi}\int\mathrm{d}\mu\;\hat\varrho(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x^\star(\mu) \tag{5.48}$$

and

$$\lim \prod_{\mu} \Big[Z(\mu_\varepsilon|\mathbf{A}^\star_{\ell,\mu})^{n_\mu} \overline{Z(\mu_\varepsilon|\mathbf{A}^\star_{\ell,\mu})}^{m_\mu} \Big] \;=\; \mathrm{e}^{\frac{2}{\pi}\int \mathrm{d}\mu\ \hat\varrho(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log Z(\mu_\varepsilon|\mathbf{A}^\star_{\ell,\mu})}. \tag{5.49}$$

This gives us

$$\begin{aligned}\phi[\hat\varrho] \;&=\; \int \mathrm{d}\mu\ \hat\varrho(\mu)\Big\{ \frac{1}{2\pi}\frac{\mathrm{d}}{\mathrm{d}\mu}\ \mathrm{Im}\Big[(q-2)\log\Big(x^\star(\mu)-\frac{1}{x^\star(\mu)}\Big) - q\log x^\star(\mu)\Big]\Big\} \\ &\quad + \frac{1}{N}\sum_{\ell=3}^{\infty}\frac{(q-1)^\ell}{2\ell}\ \mathrm{e}^{\frac{1}{\pi}\int \mathrm{d}\mu\ \hat\varrho(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\left[\ell\log x^\star(\mu)+2\log Z(\mu_\varepsilon|\mathbf{A}^\star_{\ell,\mu})\right]}.\end{aligned} \tag{5.50}$$

Since $\delta\phi[\hat\varrho]/\delta\hat\varrho(\mu) = \langle\varrho(\mu|\mathbf{A})\rangle$, the first line of (5.50) is the generator of the asymptotic spectrum in the limit $N\to\infty$, whereas the second line will give us the $\mathcal{O}(1/N)$ finite size corrections to the spectrum. In C.3 we show that, upon taking the limit $\varepsilon\to 0$, the factor inside the curly brackets in the first line indeed works out to be exactly the Kesten-McKay law (KM) [46, 67] for random regular graphs:

$$\varrho_0(\mu) \;=\; \frac{q}{2\pi}\frac{\sqrt{4(q-1)-\mu^2}}{q^2-\mu^2}\ \theta\big[2\sqrt{q-1}-|\mu|\,\big]. \tag{5.51}$$

This shows that, for regular graphs, the deformation of the measure in the ensemble (5.6) does not alter the resulting spectrum in leading order, but in sub-leading order $\mathcal{O}(N^{-1})$. In regular graphs, the Lagrange parameter $\hat\varrho(\mu)$ apparently needs to be rescaled further with N to induce a spectrum that in leading order differs from (5.51), similar to what was found in [27].

Having simplified the first line of (5.50) to $\int \mathrm{d}\mu\ \hat\varrho(\mu)\varrho_0(\mu)$, we now work out further the exponent in the second line of (5.50). First, defining $h(\mu) = -\frac{1}{\pi}\mathrm{Im}\log x^\star(\mu)$, in C.3 we show that

$$\frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu) \;=\; -\frac{1}{\pi}\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x^\star(\mu) = \frac{1}{\pi}\frac{\theta\left(2\sqrt{q-1}-|\mu|\right)}{\sqrt{4(q-1)-\mu^2}}. \tag{5.52}$$

We can evaluate the second term in the exponent using the eigenvalues $\lambda_k = 2\cos(2\pi k/\ell)$ of the adjacency matrix $A_{ij} = \delta_{i,j+1} + \delta_{i,j-1}\ (\mathrm{mod}\ \ell)$ of a length-ℓ loop, and the identity

$\mathrm{d}x^\star(\mu)/\mathrm{d}\mu = -\mathrm{i}x^\star(\mu)/\sqrt{4(q-1)-\mu^2}$:

$$\begin{aligned}
g_\ell(\mu) &= \lim_{\varepsilon\to 0}\frac{2}{\ell\pi}\mathrm{Im}\log Z(\mu_\varepsilon|\mathbf{A}^\star_{\ell,\mu}) \\
\frac{\mathrm{d}}{\mathrm{d}\mu}g_\ell(\mu) &= \lim_{\varepsilon\to 0}\frac{2}{\ell\pi}\mathrm{Im}\Big\{\frac{\mathrm{d}}{\mathrm{d}\mu}\log Z(\mu_\varepsilon|\mathbf{A}^\star_{\ell,\mu})\Big\} \\
&= \frac{2}{\ell\pi}\mathrm{Im}\Big\{\frac{\mathrm{d}}{\mathrm{d}\mu}\log\Big[\prod_{k=1}^{\ell}\sqrt{\frac{2\pi}{\mathrm{i}\Big(\lambda_k+\frac{2-q}{x^\star(\mu)}-\mu\Big)}}\Big]\Big\} \\
&= \frac{1}{\ell\pi}\sum_{k=1}^{\ell}\mathrm{Im}\Big\{\frac{1-(2-q)\frac{\mathrm{d}}{\mathrm{d}\mu}(x^\star(\mu))^{-1}}{\lambda_k+\frac{2-q}{x^\star(\mu)}-\mu}\Big\} \\
&= \frac{1}{\ell\pi}\mathrm{Im}\Big\{\sum_{k=1}^{\ell}\frac{1+\mathrm{i}\frac{q-2}{x^\star(\mu)}[4(q-1)-\mu^2]^{-\frac{1}{2}}}{2\cos(2\pi k/\ell)-\frac{q-2}{x^\star(\mu)}-\mu}\Big\}. \qquad (5.53)
\end{aligned}$$

With the above simplifications we can write both the leading two orders in N of the generating function $\phi[\hat{\varrho}]$ and of the resulting average spectrum $\varrho(\mu) = \delta\phi[\hat{\varrho}]/\delta\hat{\varrho}(\mu)$ for our ensemble (5.3), for Lagrange parameters $\hat{\varrho}(\mu) = \mathcal{O}(1)$, in the following transparent form, which represents one of the main results of this chapter:

$$\phi[\hat{\varrho}] = \int \mathrm{d}\mu\ \hat{\varrho}(\mu)\varrho_0(\mu) + \frac{1}{N}\sum_{\ell=3}^{\infty}\frac{(q-1)^\ell}{2\ell}\ \mathrm{e}^{\ell\int\mathrm{d}\mu\ \hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu)-h(\mu)]} + o(\frac{1}{N}), \qquad (5.54)$$

$$\varrho(\mu) = \varrho_0(\mu) + \frac{1}{N}\sum_{\ell=3}^{\infty}\frac{(q-1)^\ell}{2\ell}\mathrm{e}^{\ell\int\mathrm{d}\mu\ \hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu)-h(\mu)]}(\ell\frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu)-h(\mu)]) + o(\frac{1}{N}). \qquad (5.55)$$

The first and leading order term ϱ_0 corresponds to the Kesten-McKay law, as was explained before. The second term is the so called *loop series*. It is a sum over loop lengths ℓ (hence it starts at $\ell = 3$), and it need not always be convergent, but this has been shown not to necessarily pose problems [63, 64, 68]. The coefficient $(q-1)^\ell/2\ell$ in each term of the series is exactly the asymptotic expected number of loops of *finite* length ℓ inside a random regular graph [69]. It is amazing that this number can be recovered with a replica calculation just by carefully calculating the sub-leading order in $1/N$, and leads to an intuitive interpretation of (5.55). In our present ensemble (5.3), the number of loops of a given length is given by the usual number found in random regular graphs, multiplied by a factor that depends on the spectral Lagrange parameter $\hat{\varrho}(\mu)$, being $\exp(\ell\int\mathrm{d}\mu\ \hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu)-h(\mu)])$. This means that the number of loops of each length ℓ can be either increased or decreased depending of the choice of $\hat{\varrho}$. The correction to the Kesten-McKay spectrum formula due to the appearance of a single loop of length ℓ

will be given by $\ell\frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu)-h(\mu)]$, as discussed in [63, 64]. Given that the effect of the loops is additive in (5.55), we must expect that these spectral corrections come from isolated and well separated loops in the graph. In sections 4.2 and 4.4 we verify this interpretation with numerical simulations. When tuning the number of loops of length ℓ individually, we observe precisely that the plot of $\ell\frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu)-h(\mu)]$ scales as described in (5.55). The case where many loops appear that are not isolated from each other is discussed in section 4.3, here indeed it is necessary to add more terms. We wish to point out that, while a cavity approach could account for the presence of loops, it would not be able to provide information on their average number in an ensemble such as (5.3). The imaginary replica approach presented here, in contrast, has simultaneously provided for the ensemble (5.3) both the spectrum formula *and* the expected number of loops.

5.2.4 Remaining integrals over eigenvalues

In our present theory we have an as yet arbitrary functional Lagrange parameter $\hat{\varrho}(\mu)$ which controls the dependence of the graph probabilities on their expected spectra. In (5.55) we still have integrals over $\hat{\varrho}(\mu)$, of the form:

$$\mathcal{J}_\ell[\hat{\varrho}] \;=\; \int \mathrm{d}\mu\; \hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu)-h(\mu)]. \tag{5.56}$$

While expressions (5.52,5.53) for $h(\mu)$ and $g_\ell(\mu)$ will turn out useful in establishing links with previous research in a subsequent section, here we will continue the further evaluation of $\mathcal{J}_\ell[\hat{\varrho}]$ using the earlier forms

$$h(\mu) \;=\; -\frac{1}{\pi}\mathrm{Im}\log x^\star(\mu), \tag{5.57}$$

$$g_\ell(\mu) \;=\; -\frac{1}{\ell\pi}\sum_{k=1}^{\ell}\mathrm{Im}\log\Big[\mathrm{i}\Big(\cos(2\pi k/\ell)+\frac{2-q}{x^\star(\mu)}-\mu\Big)\Big]. \tag{5.58}$$

These give

$$\begin{aligned}\mathcal{J}_\ell[\hat{\varrho}] \;&=\; -\mathrm{Im}\int \mathrm{d}\mu\; \hat{\varrho}(\mu)\frac{1}{\ell\pi}\sum_{k=1}^{\ell}\frac{\mathrm{d}}{\mathrm{d}\mu}\Big[\log\Big(\cos(2\pi k/\ell)+\frac{2-q}{x^\star(\mu)}-\mu\Big)-\log x^\star(\mu)\Big]\\ &=\; \mathrm{Im}\int \mathrm{d}\mu\; \hat{\varrho}(\mu)\frac{2}{\ell\pi}\sum_{k=1}^{\ell}\frac{\mathrm{d}}{\mathrm{d}\mu}\log\Big[\frac{1}{x^\star(\mu)}\Big(\cos(2\pi k/\ell)+\frac{2-q}{x^\star(\mu)}-\mu\Big)\Big]^{-\frac{1}{2}}.\end{aligned} \tag{5.59}$$

We now use $x^\star(\mu) + x^\star(\mu)^{-1} = -\mu - (q-2)/x^\star(\mu)$, which follows directly from (5.37):

$$\mathcal{J}_\ell[\hat{\varrho}] \;=\; \mathrm{Im}\int \mathrm{d}\mu\; \hat{\varrho}(\mu)\frac{2}{\ell\pi}\sum_{k=1}^{\ell}\frac{\mathrm{d}}{\mathrm{d}\mu}\log\Big(1+\frac{\cos(2\pi k/\ell)}{x^\star(\mu)}+\frac{1}{(x^\star(\mu))^2}\Big)^{-\frac{1}{2}}. \tag{5.60}$$

In this expression we recognize the generating function of the Chebyshev polynomials $T_n(t)$ [70]. These are defined for $t \in [-1, 1]$, and can be written in explicit form as

$$T_n(t) \;=\; \mathrm{Re}\left(t+\mathrm{i}\sqrt{1-t^2}\right)^n. \tag{5.61}$$

They obey the orthogonality relation

$$\frac{2}{\pi}\int_{-1}^{1}\frac{\mathrm{d}t}{\sqrt{1-t^2}}\;T_n(t)T_m(t) \;=\; \delta_{nm}(1+\delta_{m0}) \tag{5.62}$$

as well as

$$T_n(\cos(\theta)) = \cos(n\theta), \qquad T_n(-t) = (-1)^n T_n(t). \tag{5.63}$$

The first five Chebyshev polynomials are [70]:

$$\begin{aligned} &T_0(t) = 1, && T_1(t) = t, && T_2(t) = 2t^2 - 1,\\ &T_3(t) = 4t^3 - 3t, && T_4(t) = 8t^4 - 8t^2 + 1. \end{aligned} \tag{5.64}$$

For the evaluation of (5.60), in particular, we may apply the generating function identity

$$\sum_{n=1}^{\infty} T_n(t)\frac{x^n}{n} \;=\; \log\left(1-2tx+x^2\right)^{-\frac{1}{2}} \qquad \text{for } |x|<1 \tag{5.65}$$

to the choices $x = 1/x^\star(\mu)$ and $t = -\cos(2\pi k/\ell)$, in order to obtain

$$\begin{aligned} \mathcal{J}_\ell[\hat{\varrho}] \;&=\; \mathrm{Im}\int \mathrm{d}\mu\; \hat{\varrho}(\mu)\frac{2}{\ell\pi}\sum_{k=1}^{\ell}\sum_{n=1}^{\infty}\frac{(-1)^n}{n}T_n(\cos(2\pi k/\ell))\frac{\mathrm{d}}{\mathrm{d}\mu}(x^\star(\mu))^{-n}\\ &=\; \mathrm{Im}\int \mathrm{d}\mu\; \hat{\varrho}(\mu)\frac{2}{\ell\pi}\sum_{k=1}^{\ell}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{T_n(\cos(2\pi k/\ell))}{(x^\star(\mu))^{n+1}}\frac{\mathrm{d}}{\mathrm{d}\mu}x^\star(\mu). \end{aligned} \tag{5.66}$$

We next use $\mathrm{d}x^\star(\mu)/\mathrm{d}\mu = -\mathrm{i}x^\star(\mu)/\sqrt{4(q-1)-\mu^2}$ and the short-hands

$$\begin{aligned} d_{n,\ell} &= \frac{1}{\ell}\sum_{k=1}^{\ell} T_n(\cos(2\pi k/\ell)) \\ &= \frac{1}{\ell}\sum_{k=1}^{\ell} \cos(2\pi nk/\ell) = \sum_{p\in\mathbb{Z}} \delta_{n,p\ell}. \end{aligned} \tag{5.67}$$

This results in

$$\begin{aligned} \mathcal{J}_\ell[\hat{\varrho}] &= \mathrm{Im}\int \mathrm{d}\mu\ \hat{\varrho}(\mu)\frac{2}{\pi}\sum_{n=1}^{\infty}(-1)^n\frac{d_{n,\ell}}{(x^\star(\mu))^n}\frac{\mathrm{i}}{\sqrt{4(q-1)-\mu^2}} \\ &= \sum_{n=1}^{\infty}(-1)^n d_{n,\ell}\ \frac{2}{\pi}\int \mathrm{d}\mu\ \hat{\varrho}(\mu)\mathrm{Re}\Big[\frac{1}{(x^\star(\mu))^n}\frac{1}{\sqrt{4(q-1)-\mu^2}}\Big] \\ &= \sum_{n=1}^{\infty}(-1)^n d_{n,\ell}\ \frac{2}{\pi}\int \mathrm{d}\mu\ \frac{\hat{\varrho}(\mu)}{|x^\star(\mu)|^{2n}}\mathrm{Re}\Big[\frac{\overline{x^\star(\mu)}^n}{\sqrt{4(q-1)-\mu^2}}\Big]. \end{aligned} \tag{5.68}$$

If $\mu^2 < 4(q-1)$ one has $|x^\star(\mu)|^2 = q-1$ and $\sqrt{4(q-1)-\mu^2} \in [0,\infty)$. For $\mu^2 > 4(q-1)$, on the other hand, we have $x^\star(\mu) \in \mathbb{R}$ and $\sqrt{4(q-1)-\mu^2}$ is purely imaginary. We also note that for $\mu^2 < 4(q-1)$ we may write

$$\begin{aligned} \mathrm{Re}\Big(\overline{x^\star(2t\sqrt{q-1})}^n\Big) &= (q-1)^{n/2}\ \mathrm{Re}\Big(-t+\mathrm{i}\sqrt{1-t^2}\Big)^n \\ &= (-1)^n(q-1)^{n/2}\ T_n(t). \end{aligned} \tag{5.69}$$

In combination these properties allow us to simplify $\mathcal{J}_\ell[\hat{\varrho}]$ to

$$\begin{aligned} \mathcal{J}_\ell[\hat{\varrho}] &= \sum_{n=1}^{\infty}\frac{(-1)^n d_{n,\ell}}{(q-1)^n}\ \frac{2}{\pi}\int_{-2\sqrt{q-1}}^{2\sqrt{q-1}}\frac{\mathrm{d}\mu\ \hat{\varrho}(\mu)\mathrm{Re}\big[\overline{x^\star(\mu)}^n\big]}{\sqrt{4(q-1)-\mu^2}} \\ &= \sum_{n=1}^{\infty}\frac{d_{n,\ell}}{(q-1)^{n/2}}\ \frac{2}{\pi}\int_{-1}^{1}\frac{\mathrm{d}t}{\sqrt{1-t^2}}\ \hat{\varrho}(2t\sqrt{q-1})T_n(t) \\ &= \sum_{p=1}^{\infty}\frac{1}{(q-1)^{p\ell/2}}\ \frac{2}{\pi}\int_{-1}^{1}\frac{\mathrm{d}t}{\sqrt{1-t^2}}\ \hat{\varrho}(2t\sqrt{q-1})T_{p\ell}(t). \end{aligned} \tag{5.70}$$

The above result shows that the Chebyshev polynomials form the natural basis in terms of which to express the functional Lagrange parameter $\hat{\varrho}(\mu)$, and is inserted into our spectrum formula (5.55) to give

$$\varrho(\mu) = \varrho_0(\mu) + \frac{1}{2N}\sum_{\ell=3}^{\infty}(q-1)^\ell \mathrm{e}^{\ell\mathcal{J}_\ell[\hat{\varrho}]}\frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu)-h(\mu)] + o(\frac{1}{N}). \tag{5.71}$$

It is instructive to work out $\mathcal{J}_\ell[\hat{\varrho}]$ for $\ell \geq 3$ and some simple choices of $\hat{\varrho}(\mu)$:

- $\hat{\varrho}(\mu) = \mu^3$:

  This choice corresponds to random regular graphs in which the number of triangles is controlled. We use $t^3 = \frac{1}{4}T_3(t) + \frac{3}{4}T_1(t)$ and the orthogonality relation (5.62):

$$\begin{aligned} \mathcal{J}_\ell[\hat{\varrho}] &= \sum_{p=1}^{\infty} \frac{2}{(q-1)^{(p\ell-3)/2}} \frac{2}{\pi} \int_{-1}^{1} \frac{\mathrm{d}t}{\sqrt{1-t^2}} \big[T_3(t)+3T_1(t)\big] T_{p\ell}(t) \\ &= 2\sum_{p=1}^{\infty} \frac{\delta_{3,p\ell} + 3\delta_{1,p\ell}}{(q-1)^{(p\ell-3)/2}} = 2\sum_{p=1}^{\infty} \delta_{3,p\ell} = 2\delta_{\ell 3}. \end{aligned} \tag{5.72}$$

- $\hat{\varrho}(\mu) = \mu^4$:

  This choice corresponds to random regular graphs in which the number of squares is controlled. We use $t^4 = \frac{1}{8}T_4(t) + \frac{1}{4}T_2(t) + \frac{3}{8}T_0(t)$ and the orthogonality (5.62):

$$\begin{aligned} \mathcal{J}_\ell[\hat{\varrho}] &= \sum_{p=1}^{\infty} \frac{2}{(q-1)^{(p\ell-4)/2}} \frac{2}{\pi} \int_{-1}^{1} \frac{\mathrm{d}t}{\sqrt{1-t^2}} \big[T_4(t) + 2T_2(t) + 3T_0(t)\big] T_{p\ell}(t) \\ &= 2\sum_{p=1}^{\infty} \frac{\delta_{4,p\ell} + 2\delta_{2,p\ell} + 3\delta_{0,p\ell}}{(q-1)^{(p\ell-4)/2}} = 2\sum_{p=1}^{\infty} \delta_{4,p\ell} = 2\delta_{\ell 4}. \end{aligned} \tag{5.73}$$

5.3 Applications of the general theory

5.3.1 Recovering previous results as a test

Upon making the trivial choice $\hat{\varrho}(\mu) = 0$ we return to the conventional ensembles with uniform probabilities, and our equations (5.54,5.55) recover the natural spectrum fluctuations of random regular graphs, as previously studied in detail in [66] and with the traditional replica method (where $n \to 0$) in [63]:

$$\varrho(\mu) = \varrho_0(\mu) + N^{-1}\varrho_1(\mu) + o(N^{-1}) \tag{5.74}$$

with

$$\varrho_1(\mu) = \sum_{\ell=3}^{\infty} \frac{(q-1)^\ell}{2} \frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu) - h(\mu)]. \tag{5.75}$$

This series was summed in [63], and we can connect the result of the summation, in the notation of [63], directly to the theory developed in the present chapter as follows:

$$\begin{aligned}\varrho_1(\mu) &= h(\mu)\ \mathrm{Re}\left[\frac{(q-1)g_c(\mu)}{1-(q-1)g_c(\mu)}+\frac{(q-1)g_c^2(\mu)}{1-(q-1)g_c^2(\mu)}\right]\\ &\quad +h(\mu)\ \mathrm{Re}\left[\sum_{\ell=3}^{\infty}(q-1)^\ell\frac{g_c^{3\ell}(\mu)}{1-g_c^\ell(\mu)}-K(g_c(\mu))\right],\end{aligned} \tag{5.76}$$

in which now

$$g_c(\mu) = -1/x^\star(\mu), \tag{5.77}$$

$$K(g) = (q-1)g+q(q-1)g^2+(q-1)^2g^4. \tag{5.78}$$

Here $h(\mu)$ and $x^\star(\mu)$ are given in (5.52) and (5.38), respectively.

As a second test we can make the special choices $q=2$ and $\hat{\varrho}(\mu)=\alpha\mu^3$, resulting in the ensemble that was studied in Chapter 2, [27], via direct combinatorics, i.e. without the replica method. This particular model represents the simplest solvable non-uniform random graph ensemble with tuneability of the frequency of short loops. First, by setting $q=2$ our general results (5.52,5.53,5.55) simplify greatly. We now find that

$$\varrho_0(\mu) = \frac{1}{\pi}\frac{\theta(2-|\mu|)}{\sqrt{4-\mu^2}}, \qquad \frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu)=\varrho_0(\mu) \tag{5.79}$$

$$\frac{\mathrm{d}}{\mathrm{d}\mu}g_\ell(\mu) = \lim_{\varepsilon\to 0}\frac{1}{\ell\pi}\mathrm{Im}\Big\{\sum_{k=1}^{\ell}\frac{1}{2\cos(2\pi k/\ell)-\mu-\mathrm{i}\varepsilon}\Big\} = \frac{1}{\ell}\sum_{k=1}^{\ell}\delta\Big[\mu-2\cos(\frac{2\pi k}{\ell})\Big] \tag{5.80}$$

Upon inserting also $\hat{\varrho}(\mu)=\alpha\mu^3$ into (5.55) we need the values of

$$\int \mathrm{d}\mu\ \mu^3\varrho_0(\mu) = 0 \tag{5.81}$$

$$\int \mathrm{d}\mu\ \mu^3\frac{\mathrm{d}}{\mathrm{d}\mu}g_\ell(\mu) = \frac{8}{\ell}\sum_{k=1}^{\ell}\cos^3(\frac{2\pi k}{\ell}) = 2\delta_{\ell 3}+8\delta_{\ell 1} \tag{5.82}$$

With this the ensemble spectrum becomes

$$\varrho(\mu) = \varrho_0(\mu)+\frac{1}{2N}\sum_{\ell=3}^{\infty}\mathrm{e}^{6\alpha\delta_{\ell 3}}\Big\{\frac{1}{\ell}\sum_{k=1}^{\ell}\delta\Big[\mu-2\cos(\frac{2\pi k}{\ell})\Big]-\frac{1}{\pi}\frac{\theta(2-|\mu|)}{\sqrt{4-\mu^2}}\Big\}+o(\frac{1}{N}) \tag{5.83}$$

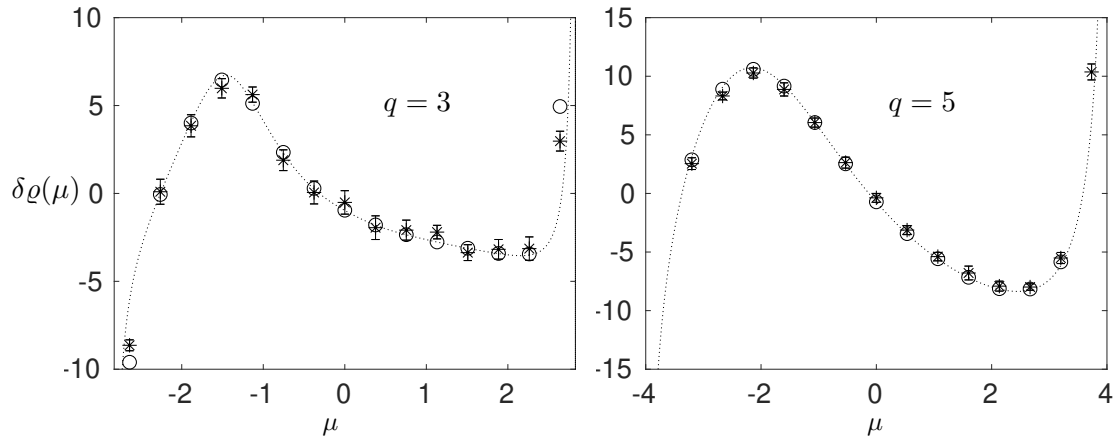


Figure 5.1: Average spectral densities for q-regular graphs sampled from (5.6). We show the rescaled finite size deviations from the standard Kersten-McKay formula $\varrho_0(\mu)$, by plotting $\delta\varrho(\mu) = N[\varrho(\mu) - \varrho_0(\mu)] = \varrho_1(\mu) + \tilde{\varrho}_1(\mu)$. Left panel: $q = 3$, $N = 1000$ and $\alpha = 0.416$, giving average clustering coefficient $\langle C(\mathbf{A})\rangle = 0.016$. Right figure: $q = 5$, $N = 2000$ and $\alpha = 0.431$, giving average clustering coefficient $\langle C(\mathbf{A})\rangle = 0.02$. Each marker shows the average spectral density contribution obtained from 200 histograms of samples of (5.6), generated with an appropriate MCMC process, and error bars indicate $\pm$ one standard deviation. The dotted line shows the theoretical prediction (5.85,5.87), and circles show the density prediction computed for exactly the eigenvalue bins that were also used for the histograms of the simulation samples.

and for the triangle density $m(\alpha) = \int \mathrm{d}\mu\ \varrho(\mu)\mu^3$ we obtain

$$m(\alpha) = N^{-1}\mathrm{e}^{6\alpha} \tag{5.84}$$

These results are indeed identical to those derived combinatorially in [27].

5.3.2 Triangularly constrained regular graph ensemble with arbitrary degree

We proceed to apply the general theory developed in the previous section to the graph ensemble (5.6) with controlled numbers of triangles, i.e. with $\hat{\varrho}(\mu) = \alpha\mu^3$, but now for arbitrary values of the degree q where the direct combinatorial approach of [27] is no longer possible. We can start directly by inserting (5.72) into (5.71), upon adding the control

parameter α, giving

$$\begin{aligned}\varrho(\mu) &= \varrho_0(\mu)+\frac{1}{2N}\sum_{\ell=3}^{\infty}(q-1)^{\ell}\mathrm{e}^{6\alpha\delta_{\ell 3}}\frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu)-h(\mu)]+o(\frac{1}{N})\\ &= \varrho_0(\mu)+\frac{1}{N}\varrho_1(\mu)+\frac{1}{N}\tilde{\varrho}_1(\mu)+o(\frac{1}{N}), \qquad (5.85)\\ \tilde{\varrho}_1(\mu) &= \frac{1}{2}(q-1)^3(\mathrm{e}^{6\alpha}-1)\frac{\mathrm{d}}{\mathrm{d}\mu}[g_3(\mu)-h(\mu)]. \qquad (5.86)\end{aligned}$$

Here $\varrho_1(\mu)$ is the function (5.75) that already appeared in the spectrum of the non-deformed ensembles of [63], and for which we can use the resummation (5.76). The impact of controlling the graph probabilities (5.6) with a nonzero Lagrange parameter $\hat{\varrho}(\mu)=\alpha\mu^3$ is fully concentrated in $\tilde{\varrho}_1(\mu)$. We next insert our earlier expressions for $g_\ell(\mu)$ and $h(\mu)$ into (5.86) and simplify the result where possible:

$$\begin{aligned}\tilde{\varrho}_1(\mu) =& \frac{(q-1)^3}{2\pi}(\mathrm{e}^{6\alpha}-1)\;\theta[2\sqrt{q-1}-|\mu|]\\ &\times\Big\{\mathrm{Im}\Big[\frac{1}{3}\sum_{k=1}^{3}\frac{1+\mathrm{i}\frac{q-2}{x^\star(\mu)}[4(q-1)-\mu^2]^{-\frac{1}{2}}}{2\cos(2\pi k/3)-\frac{q-2}{x^\star(\mu)}-\mu}\Big]-\frac{1}{\sqrt{4(q-1)-\mu^2}}\Big\}\\ =& \frac{(q-1)^3}{2\pi}(\mathrm{e}^{6\alpha}-1)\frac{\theta[2\sqrt{q-1}-|\mu|]}{\sqrt{4(q-1)-\mu^2}}\Big\{\frac{q-2}{3}\Big[\frac{2q+\mu}{q^2-3(q-1)+\mu q+\mu^2}+\frac{1}{q-\mu}\Big]-1\Big\}. \qquad (5.87)\end{aligned}$$

The results of testing this prediction against numerical simulations are shown in Figure 5.1, and reveal excellent agreement. In the simulations we sampled numerically from (5.6) with an edge swap based Markov Chain Monte Carlo algorithm (MCMC) with nontrivial acceptance probabilities. Edge swaps are accepted or rejected depending on the change in the number of loops and on the change in the possible number of possible swaps. This corrects for entropic effects, see e.g. [9] or [6]. Since we work with a system of finite size N, our predictions refer to the *average* eigenvalue density, not to the density of individual graph instances. The error bars in Figure 5.1 are computed following 10 different initialization seeds of the MCMC algorithm, consisting of distinct regular graphs sampled uniformly with a configuration model algorithm. Following each initialization, 20 samples were taken, separated in algorithmic time by $\sim 10^3$ accepted MCMC swaps per link in the graph.

We can also calculate the expected triangle density $m(\alpha)=\int\mathrm{d}\mu\;\varrho(\mu)\mu^3$ for the ensemble (5.6). It is easier to do this by integrating over (5.71) rather than via (5.85,5.87),

although both routes give the same result:

$$\begin{aligned} m(\alpha) &= \frac{1}{2N}\sum_{\ell=3}^{\infty}(q-1)^{\ell}\mathrm{e}^{\alpha\ell\int \mathrm{d}\mu\ \mu^3\frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu)-h(\mu)]}\int \mathrm{d}\mu\ \mu^3\frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu)-h(\mu)]+o(\frac{1}{N}) \\ &= \frac{1}{2N}\sum_{\ell=3}^{\infty}(q-1)^{\ell}\mathrm{e}^{2\delta_{\ell 3}\alpha\ell}2\delta_{\ell 3}+o(\frac{1}{N}) \\ &= \frac{1}{N}(q-1)^3\mathrm{e}^{6\alpha}+o(\frac{1}{N}). \end{aligned} \tag{5.88}$$

This formula gives very accurate results for α values up to a certain point, defined as $\alpha_1(N)$ in the next section. This can be seen very clearly in figures 5.2 and 5.4, where we test its predictions for ensembles (5.6) with $q = 3$. Since $m(\alpha)$ represents an ensemble average, we compare (5.88) against the average loop density over multiple graphs drawn from the ensemble, $\hat{m}(\alpha) = M^{-1}\sum_{m=1}^{M}\mathrm{Tr}(\mathbf{A}_m^3)$. Figures 5.2 and 5.4 show averages and standard deviations of the estimator $\hat{m}(\alpha)$, for 50 different small samples sampled during MCMC simulations, separated in algorithmic time by $\sim 10^3$ swaps per link (in order to ensure independence of the M samples).

We have developed a theory that quantifies the $\mathcal{O}(1/N)$ effects on the eigenvalue spectrum of probability deformations in ensembles of the general family (5.3), in which loops can be induced via the functional Lagrange parameter $\hat{\varrho}(\mu)$, and we applied our results to a specific member (5.6) of this family. As mentioned before, the theory implicitly assumes that loops inside the graph are far away from each other. As the control parameter α in (5.6) is increased for fixed system size N, we must therefore expect the behaviour of the ensemble to start deviating from the predictions (5.85,5.87,5.88) as soon as the loops start to interact. This can indeed be seen in figure 5.2. As α increases the clustering coefficient starts deviating from (5.88), which is shown as a dashed line. Mathematically, one can explain the deviations from (5.88) as the emergence of higher order corrections to the saddle point approximation, $\mathcal{O}(N^{-\gamma})$ with $\gamma > 1$, that were not incorporated into the replica calculation. These would account for the presence of loops that are not isolated from each other. The accuracy of (5.85,5.87) and (5.88) suggests that calculating higher order corrections in the replica calculation would improve our predictions, but this would require of course a much more complicated calculation. In the next section we will explore what happens as we keep increasing α beyond the validity of our replica calculation.

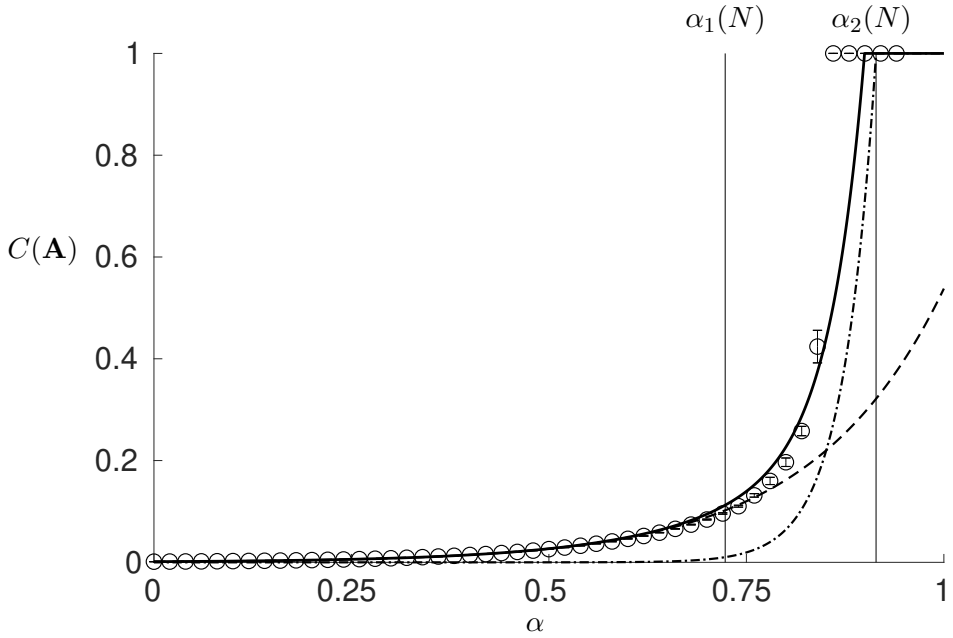


Figure 5.2: Plot of the clustering coefficient $C(\mathbf{A})$ versus α. Circles show results from MCMC sampling with $N = 1000, q = 3$ (average plus/minus one standard deviation). Solid line: predicted values computed from the theory (5.88), via $\langle C(\mathbf{A})\rangle = m(\alpha)/q(q-1)$. We also show separately the two distinct contributions to the theoretical prediction, viz. $\langle C(\mathbf{A})\rangle_T$ (those from disconnected triangles, dashed line) and $\langle C(\mathbf{A})\rangle_K$ (those from triangles in cliques, dotted dashed line). Typical graph examples generated within each α regime are shown in Figure 5.3.

5.3.3 Phases of the ensemble and the shattering transition

We will now give a qualitative picture of the behaviour of the ensemble (5.6) for all values of $\alpha \in [0, \infty)$. We will focus on $q \geq 3$, since the case $q = 2$ was already covered in Chapter 2, [27]. A similar discussion is already included in Chapter 3, nevertheless this one is specific to the regular case as the availability of more exact expressions and less possible subgraphs makes it special. In MCMC simulations one observes three distinct regimes, which are not phases in a rigorous thermodynamic sense, but size dependent ranges of α values that exhibit qualitatively different phenomenology:

- *Small* α: *connected phase*

  The triangle promoting probability bias in the ensemble introduces isolated loops embedded in the giant component. Here the analysis of the previous section should apply, as is confirmed in figures 5.2 and 5.4 for different values of q. Indeed one observes only small deviations from (5.88), as one approaches the next phase.

- *Intermediate* α: *partially connected phase*

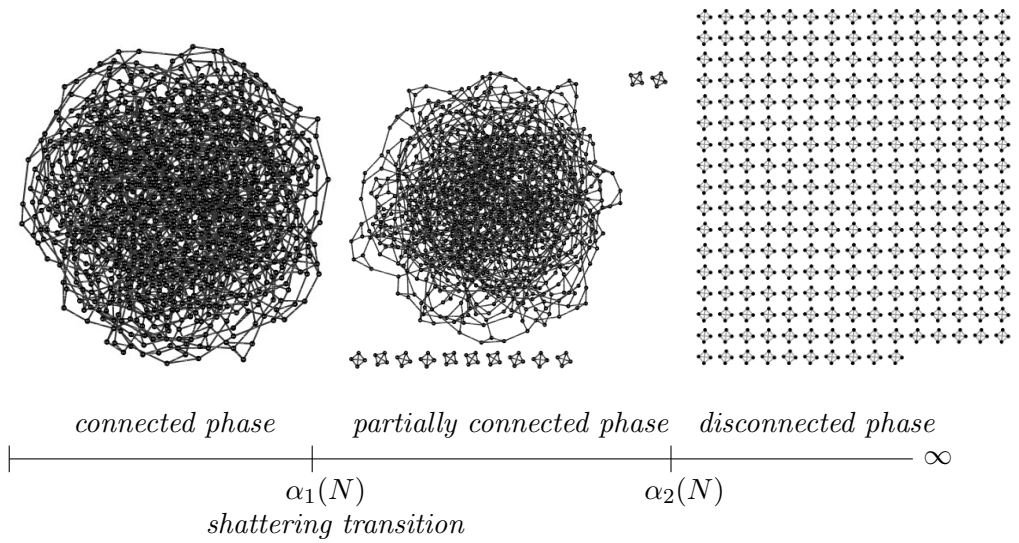


Figure 5.3: Three typical 3-regular graphs, of size $N = 1000$, sampled numerically via MCMC from the canonical ensemble (5.6). The value of the tuning parameter α increases from left to right, and each graph shown is generated from one of the three distinct phases defined in Figure 5.2.

Edges can now be part of more than one triangle, and the graphs contain an increasing number of cliques of $q+1$ nodes, denoted by K_{q+1}. The triangle density and hence the clustering coefficient grow considerably faster than in the previous phase, increasingly so for larger degrees q.

- *Large* α: *disconnected phase*

  Here the graphs break down completely into large collections of those cliques that had started to appear in the previous phase. The resulting configurations correspond to q regular graphs with the maximum possible number of triangles. In analogy with physics, we call these ground states.

We label the transition points between the phases $\alpha_1(N)$ and $\alpha_2(N)$, see figure 5.2. Both $\alpha_1(N)$ and $\alpha_2(N)$ grow logarithmically with N. We refer to the transition from connected to partially connected as the *shattering transition* to highlight its topological nature. In figure 5.3 we show typical graphs sampled via MCMC in the three phases.

In order to complement the previous distinctions with quantitative estimates, we will present the results from Chapter 3 for the specific case of regular graphs. The generating function $\phi(\alpha)$, (3.23, and the loop density $m(\alpha)$, (3.17) using an Poissonian approximation for the distribution of triangles in cliques. For regular graphs *both* expected values are

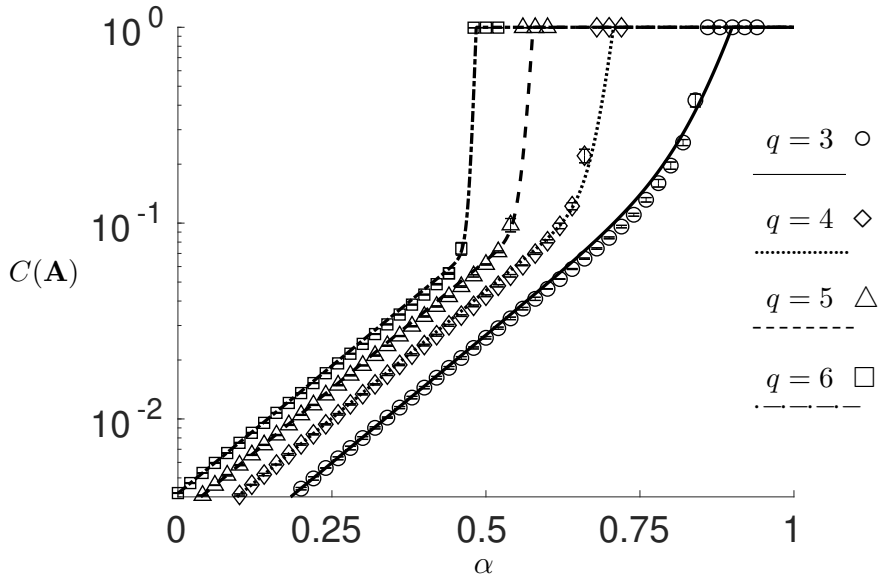


Figure 5.4: In this figure we show the agreement between the clustering coefficients predicted by (5.91), with lines, and the values measured MCMC simulations (markers, showing average plus/minus one standard deviation) with $N = 1000$. Full details on the number of samples generated and their separation in MCMC edge swaps are given in the main text. These results confirm that (5.91) captures the essence of the phenomenology of the ensemble.

known rigorously, [5, 71]. Upon approximating $N!/(N-q-1)! \approx N^{q+1}$ we obtain

$$\lambda_K = \frac{1}{N^{\frac{1}{2}(q-2)(q+1)}} \frac{(q!)^{q+1}}{q^{\frac{q}{2}(q+1)}(q+1)!}. \tag{5.89}$$

See C.4 for details.

We can then evaluate (3.23) and (3.24) exactly,

$$\phi(\alpha) \;\approx\; \frac{1}{N}\big[\lambda_t \mathrm{e}^{6\alpha} + \lambda_K \mathrm{e}^{\alpha q(q^2-1)} - \lambda_t - \lambda_K + \log \mathcal{N}_q\big]. \tag{5.90}$$

From this directly follows $m(\alpha) = \partial\phi(\alpha)/\partial\alpha$:

$$m(\alpha) \;\approx\; \frac{1}{N}(q-1)^3 \mathrm{e}^{6\alpha} + \frac{1}{N^{1+\frac{1}{2}(q-2)(q+1)}} \frac{(q!)^{q+1}}{q^{\frac{q}{2}(q+1)}(q-2)!} \mathrm{e}^{\alpha q(q^2-1)}. \tag{5.91}$$

The first term coincides with formula (5.88) from the previous section. The second term in (5.91) represents the impact of cliques, and, according to figure 5.4, accounts for most of the deviations from (5.88). In spite of our approximation of only accounting for isolated loops and isolated cliques, the resulting description is seen to give very good agreement with simulations for the whole range of α values.

As one might expect, the MCMC sampling algorithm slows down as it approaches the ground state. While we will not carry out a detailed dynamical analysis, we will mention the MCMC process slows down considerably precisely in the partially connected phase. To obtain good (i.e. sufficiently independent) samples even close to the ground state, we increased the number of accepted swaps per link in between samples beyond $\alpha_1(N)$ to values in the range of $10^4 - 10^5$ accepted swaps per link. We also increased considerably the waiting time before the first sample to $\sim 10^8$ accepted swaps per link, to allow the system to escape from possible metastable states.

Expression (5.91) has a clear interpretation: the expected number of subgraphs of the types T and K are boosted independently when increasing α, each with an exponential factor in accordance with the model (5.6). This already provides an explanation for the phases described previously. If we denote by $m_t(\alpha)$ and $m_K(\alpha)$ the first and second term of (5.91), then we can describe the phases in terms of the relation between these two terms.

- *Small α: connected phase*

  Here we have $m_t(\alpha) \gg m_K(\alpha)$. Even though cliques K_{q+1} may be present, their probability is too small to be relevant.

- *Intermediate α: partially connected phase*

  Here $m_K(\alpha)$ becomes significant. We may define the onset $\alpha_1(N)$ of the partially connected phase to be the point where $m_K(\alpha) = \eta m_t(\alpha)$ for some finite $\eta \in (0,1)$. Here we chose $\eta = 1/10$, which was found appropriate in the ranges $q \leq 6$ and $N \leq 2000$. The shattering transition point is then given by

$$\alpha_1(N) = \frac{1}{2}\frac{(q-2)(q+1)}{q(q^2-1)-6}\log N + \frac{1}{q(q^2-1)-6}\log\left(\eta\frac{(q-1)^2 q^{\frac{q}{2}(q+1)-1}}{(q!)^q}\right). \tag{5.92}$$

- *Large α: disconnected phase*

  Here the contribution from $m_K(\alpha)$ dominates, and the whole graph is made of disconnected cliques. The critical value $\alpha_2(N)$ marking the start of this phase is defined by the instance where $m(\alpha_2(N)) = q(q-1)$, i.e. where the maximum possible density of loops is achieved. We can replace $m(\alpha_2(N))$ by $m_K(\alpha_2(N))$ since the contribution from disconnected triangles is now very small, giving

$$\alpha_2(N) = \frac{1}{2}\frac{(q-2)(q+1)+2}{q(q^2-1)}\log(N) + \frac{\log(q)}{2(q-1)} - \frac{\log(q!)}{q^2-1}. \tag{5.93}$$

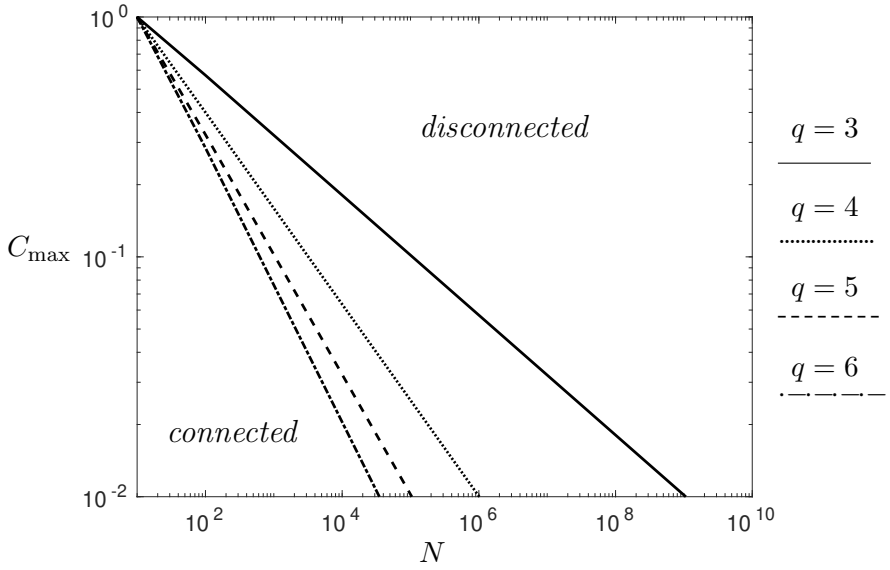


Figure 5.5: The upper bound $C_{\max}$ on the tuneable level of clustering within the giant component, for graphs from the ensemble (5.6), plotted against the graph size N. This is shown for different degrees of the regular random graphs. Values for clustering above the lines cannot be achieved in the connected phase, but would require the formation of isolated cliques.

Even though $m_K(\alpha)$ in (5.91) has a higher power of $1/N$ compared to $m_K(\alpha)$, the prefactor of α for $m_K(\alpha)$ is higher in the exponential, viz. $(q+1)q(q-2) > 6$ for $q \geq 3$. This means that for any N, the clique contribution $m_K(\alpha)$ will always grow faster with α than $m_t(\alpha)$, which implies it will always surpass it for large enough α. We do not claim that graphs will only be made either of isolated loops in a giant component, but only that knowing the behaviour of the two quantities $m_t(\alpha)$ and $m_K(\alpha)$ appears sufficient to understand the overall behaviour of the loopy ensemble (5.6).

Let us briefly discuss the potential practical utility of (5.6) in light of the previous results. Our ensemble is the maximally unbiased random graph ensemble over regular graphs that satisfies the condition of having a particular clustering coefficient C. In order to achieve one's desired value of C it is only necessary to find the appropriate $\alpha(C)$ by solving $C = C(\alpha) = m(\alpha)/q(q-1)$ using (5.91). However, if one's interest is in using (5.6) as a null model for real networks with link clustering, the presence of cliques is undesirable. If we aim to generate graphs with a single component and a nontrivial number of loops, we need to stay in the connected phase. Moreover, in this phase we have a very accurate control of $m(\alpha)$ and the spectral density through (5.85). We conclude that (5.6) can be a

useful null model when $C \in (0, C(\alpha_1(N)))$. In that clustering range we can simply take

$$\alpha(C) = \frac{1}{6}\log\left(NqC/(q-1)^2\right). \tag{5.94}$$

While the shattering transition occurs at $\alpha = \alpha_1(N)$, the finite size nature of the problem makes it possible that some cliques appear already somewhat earlier. However, it is clear from Figure 5.2 that an upper bound to the level of clustering achievable without cliques is given by $C_{\max} = C_T(\alpha_2(N))$, the contribution to clustering from disconnected triangles at the transition point to the disconnected phase. Values $C \geq C_{\max}$ are impossible to achieve in the ensemble (5.6) without triangles appearing outside the giant component, and additional constraints would have to be introduced into the model to prevent the formation of isolated cliques. This dependence on N of the hard upper bound, shown also in Figure 5.5, is somewhat unexpected. In particular,

$$\lim_{N\to\infty} C_{\max} = \lim_{N\to\infty} C_T(\alpha_2(N)) = 0. \tag{5.95}$$

Hence for very large sizes, $N \gg 1$, even very small clustering coefficients are not accessible in the connected regime. This can be understood intuitively as an entropic effect as discussed in Chapter 3, see Appendix B.

5.3.4 Other ensembles

So far we have focused specifically on the ensemble (5.6) as the simplest nontrivial instance of the more general family (5.5), suitable for testing limits and for developing further our intuition for the phenomenology of 'loopy' random graph ensembles. However, we have the more general results (5.70,5.71), applicable to any functional Lagrange parameter $\hat{\varrho}[\mu]$, with the key integral expressed as an expansion in Chebyshev polynomials. We will now turn to other choices for $\hat{\varrho}[\mu]$.

Our first choice is $\hat{\varrho}[\mu] = \alpha\mu^3 + \beta\mu^4$, which generalizes the ensemble (5.6) in that we now control the number of closed paths of both length 3 and length 4:

$$p(\mathbf{A}) = \frac{\mathrm{e}^{\alpha\mathrm{Tr}(\mathbf{A}^3)+\beta\mathrm{Tr}(\mathbf{A}^4)}}{Z(\alpha,\beta)}\prod_{i=1}^{N}\delta_{q,\sum_j A_{ij}}, \tag{5.96}$$

$$Z(\alpha,\beta) = \sum_{\mathbf{A}\in G}\mathrm{e}^{\alpha\mathrm{Tr}(\mathbf{A}^3)+\beta\mathrm{Tr}(\mathbf{A}^4)}\prod_{i=1}^{N}\delta_{q,\sum_j A_{ij}}. \tag{5.97}$$

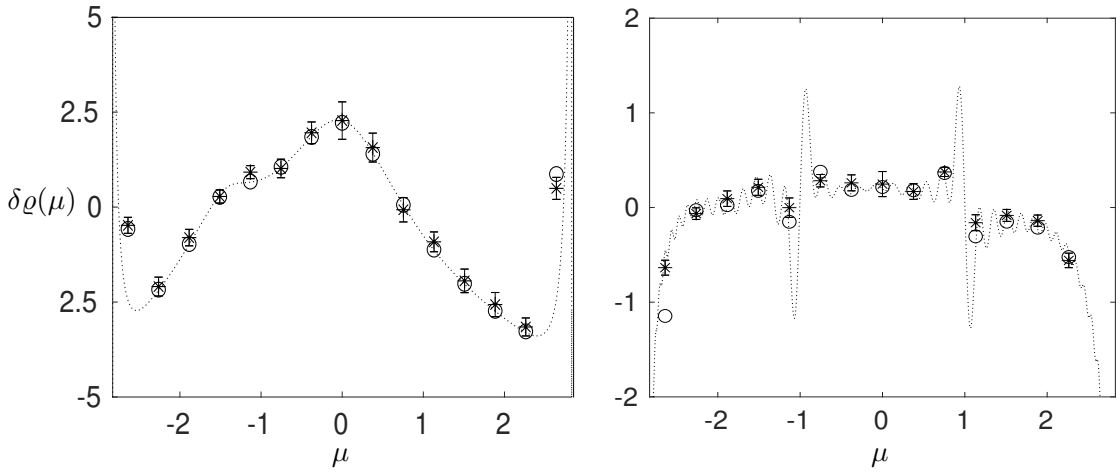


Figure 5.6: Average spectral densities for more complicated spectrally constrained q-regular graph ensembles. As before we show the rescaled finite size deviations from the Kesten-MacKay law $\delta\varrho(\mu) = N[\varrho(\mu) - \varrho_0(\mu)]$. Left: results for the ensemble (5.98), with $q = 3$, $N = 2000$, and $\alpha = \beta = 0.2$. Right: results for the ensemble (5.100), with $q = 3$, $N = 1000$, and $\alpha = 0.5$. Markers show the average density computed from from histograms of samples obtained via MCMC simulations. The dotted line shows the theoretical predictions, circles show the density prediction for the exact bins as those used for the histograms of the numerical samples. See the main text for further details.

The calculations for this ensemble (5.96) are very similar to those carried out for (5.6), which allows us to be brief. We have already computed the relevant integrals in (5.72,5.73), and we can therefore immediately proceed to the spectral density:

$$\begin{aligned}\varrho(\mu) &= \varrho_0(\mu) + \frac{1}{N}\varrho_1(\mu) + \sum_{\ell=3}^{\infty}\frac{(q-1)^\ell}{2N}\Big(\mathrm{e}^{6\alpha\delta_{3,\ell}+8\beta\delta_{4,\ell}}-1\Big)\frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu)-h(\mu)] + o(\frac{1}{N}) \\ &= \varrho_0(\mu) + \frac{1}{N}\varrho_1(\mu) + \frac{(q-1)^3}{2N}(\mathrm{e}^{6\alpha}-1)\frac{\mathrm{d}}{\mathrm{d}\mu}[g_3(\mu)-h(\mu)] \\ &\quad + \frac{(q-1)^4}{2N}(\mathrm{e}^{8\beta}-1)\frac{\mathrm{d}}{\mathrm{d}\mu}[g_4(\mu)-h(\mu)] + o(N^{-1}),\end{aligned} \tag{5.98}$$

in which the functions $g_\ell(\mu)$ and $h(\mu)$ are given in (5.57,5.58). A comparison of the predicted spectrum (5.98) with measurents in MCMC simulations, for $N = 2000$ and $\alpha = \beta = 0.2$, is shown in the left panel of Figure 5.6. The MCMC algorithm used was similar to the one described before, but now they also require monitoring the evolution of $\mathrm{Tr}(\mathbf{A}^4)$ (in addition to $\mathrm{Tr}(\mathbf{A}^3)$), as both appear in the move acceptance probabilities. In each run 100 samples were generated from each initial seed, after a burn-in (waiting time) of $\sim 10^3$ swaps per link. Error bars give the standard deviation corresponding to fluctuations between 10 different initial seeds, so that a total of 1000 graphs were averaged. As in the previous case, we recover the results from [27] when setting $q = 2$. As one would

expect, (5.98) is only valid in the vicinity of $(\alpha,\beta)=(0,0)$, to avoid the emergence of extensively many small fully connected q-regular bipartite graphlets, which maximize the number of 4−loops around a node.

Our second alternative choice for $\hat{\varrho}[\mu]$ is the following block function, which introduces a bias in the graph probabilities depending on the number of eigenvalues inside the interval $[-1,1]$:

$$\hat{\varrho}(\mu) \quad = \alpha\ \theta(1-|\mu|). \tag{5.99}$$

Now we have $N\int \mathrm{d}\mu\ \hat{\varrho}(\mu)\varrho(\mu|\mathbf{A}) = \alpha\mathcal{I}(\mathbf{A}|[-1,1])$, where $\mathcal{I}(\mathbf{A}|[-1,1])$ denotes the number of eigenvalues of $\mathbf{A}$ inside the interval $[-1,1]$. In contrast to powers of μ, understanding intuitively the topological effects of the choice (5.99) is not straightforward, notwithstanding the clear nontrivial effect on the observed spectrum. In this case we have $\mathcal{J}_\ell[\hat{\varrho}] = \alpha\int \mathrm{d}\mu\ \theta[1-|\mu|]\frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu)-h(\mu)] \neq 0$ for all ℓ, so we introduce a (sufficiently large) cutoff L in the summation of (5.71). Since $|\mathcal{J}_\ell[\hat{\varrho}]| \ll 1$ we set this integral to zero for $\ell > L$, as was done previously in [64], leaving the truncation

$$\begin{aligned}\varrho(\mu) \quad &= \quad \varrho_0(\mu) + \frac{1}{N}\varrho_1(\mu) \\ &\quad + \sum_{\ell=3}^{L} \frac{(q-1)^\ell}{2N}\Big[\mathrm{e}^{\alpha\ell\int_{-1}^{1}\mathrm{d}\mu\ \frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu)-h(\mu)]}-1\Big]\frac{\mathrm{d}}{\mathrm{d}\mu}[g_\ell(\mu)-h(\mu)],\end{aligned} \tag{5.100}$$

in which the integrals can be worked out in more explicit form, as we did for the previous cases. In figure 5.6 (right panel) we compare the prediction (5.100) with results from numerical MCMC samples, and observe a good agreement. We point out that generating graph samples from the spectrally constrained ensemble (5.5, 5.99) numerically is considerably more computationally expensive than for the previous models. Here, instead of the number of triangles or squares, the number of eigenvalues inside the interval $[-1,1]$ has to be monitored. This requires that the full set of eigenvalues of the graph $\mathbf{A}$ has to be calculated after *each* edge swap, which necessitated parallel execution in multi-core computers, to reduce CPU time to a few weeks. We seen in Figure 5.6 that the deviations from $\varrho_0(\mu)$ are quite small, nevertheless they are nontrivial and are predicted accurately. To measure spectra at this level of detail, we averaged over 10^4 graphs, separated during the MCMC process by ~ 1 swaps per link. Error bars are obtained by splitting this data set in groups of 10.

The above results are similar in form to the ones derived for weighted graphs in [58], the main difference is that in [58] a second set of replicas with the traditional limit $n \to 0$ is introduced to get the spectrum. It is interesting to see that with the functional formalism (5.5) both the observable $\int \mathrm{d}\mu\ \hat{\varrho}(\mu)\varrho(\mu)$ and the spectrum $\varrho(\mu)$ itself are calculated at the same time.

5.4 Discussion

In this chapter we have extended and applied an analytic approach for describing constrained maximum entropy ensembles of finitely connected random loopy graphs of large but finite size. We focused on regular random graphs with soft constrained adjacency matrix eigenvalue spectra. We were able to develop a general theory describing the $\mathcal{O}(1/N)$ contributions to the expected eigenvalue spectrum, through the use of an infinite number of replica indices taking values in the imaginary axis [47], and building on techniques from earlier studies such as [39, 58, 63, 64].

The simplest nontrivial spectrally constrained ensembles are those in which the spectral constraint reduces to a soft constraint on the number of triangles. We quantified the behaviour of such systems, which following [27] we have come to regard as the 'harmonic oscillators' of loopy graph ensembles, and showed how they allow for fine tuning of their average clustering coefficients. A limitation on their use as null models for regular graphs with nontrivial clustering is that there is a maximum achievable clustering coefficient, whose value depends on the size of the graph, beyond which the ensemble undergoes a transition into a new phase, where high clustering levels are achieved by the graph fracturing into extensively many disconnected cliques. We presented a precise analytic estimate for an upper bound on the maximum clustering coefficient that is achievable without fracturing of the graph. We also showed how the general theory applies to other spectrally constrained ensembles, such as those where both the number of triangles and the number of squares are controlled, and to ensembles where the spectral constraint reduces to a count of the number of adjacency matrix eigenvalues in a given interval. We carried out numerical simulations via MCMC processes based on edge swaps with nontrivial acceptance probabilities, which are themselves generally nontrivial in view of the need to recompute eigenvalue spectra after each accepted move. In comparing triangle counts and spectra, we found excellent agreement between the theoretical predictions and

the MCMC measurements in all cases, provided we remain in the parameter regime where higher orders in N of the generating function are not yet important.

The most natural generalization of the presently studied family of models would be to extend the imaginary replica approach to sparse graphs with an arbitrary degree distribution $p(k)$. This is done in the next chapter. In addition, controlling larger finite loop lengths, in the same way as in [27], could be explored. As an example we could set $\hat{\varrho}(\mu) = \alpha\mu^5$, which would introduce a bias for the closed paths of length 5. The spectral calculation upon making this modification would only be a matter of calculating $\mathcal{J}_5[\hat{\varrho}]$. Our analysis in section 4.3 relies on splitting the number of loops among different possible subgraphs favoured by the bias. Since the statistical properties of these subgraphs in uniform RRG's is detailed in [71], it is possible to develop a similar approximation as in (5.90). The range of validity of this approximation is yet to be tested numerically, but we anticipate a quick transition into a disconnected phase. Work in both this direction is currently being explored by the author.

In addition, it would be interesting to explore further the possibility of controlling short loops in finitely connected graphs without this being realized microscopically by such graphs fracturing into extensively many disconnected graphlets, even at high loop densities. This would seem to require more complicated choices of the functional Lagrange parameter $\hat{\varrho}(\mu)$ than the ones studied so far, possibly including choices that scale differently with N. Also further research in this direction is currently being explored by the author.

Chapter 6

Imaginary replicas for graphs with arbitrary degree distribution

In this chapter, we present the extension of the previous case of regular graphs to the case of graphs with arbitrary degree distribution. We assume the degree sequence of the graph $\{k_i\}_{i=1,\dots,N}$ is a sample of a given degree distribution $p(k)$. We will only require the first and second moments of this distribution to be finite to develop our theory.

In the previous chapter the solution of the problem was reduced to solving the distributional equation for $W(\mathbf{X})$, (5.33). In that case the solution was a δ-distribution, $W(\mathbf{X}) = \delta(\mathbf{X} - \mathbf{X}^*)$, where $\mathbf{X}^*$ could be written out explicitly in terms of μ_ε. For an arbitrary degree distribution the solution is typically much more complicated, and to this day no analytical solutions have been found for this type of equation, nevertheless it is possible to approach it numerically with a variation of the population dynamics algorithm. We show how we can recover previously developed theory for the uniform case and also get nontrivial results for $\hat{\varrho}(\mu) = \alpha\mu^3$.

6.1 The model

The spectrally constrained configuration model is defined by

$$\begin{aligned} p(\mathbf{A}) =& \frac{1}{Z[\hat{\varrho}]} \mathrm{e}^{N\int \mathrm{d}\mu\, \hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})} \prod_{i=1}^{N} \delta_{k_i,\sum_j A_{ij}}, \\ Z[\hat{\varrho}] =& \sum_{\mathbf{A}} \mathrm{e}^{N\int \mathrm{d}\mu\, \hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})} \prod_{i=1}^{N} \delta_{k_i,\sum_j A_{ij}}. \end{aligned} \tag{6.1}$$

The objective as usual is to calculate the generating function,

$$\begin{aligned}\phi[\hat{\varrho}] =& \frac{1}{N}\log\sum_{\mathbf{A}} \mathrm{e}^{N\mathrm{d}\mu\hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})}\prod_{i=1}^{N}\delta_{k_i,\sum_j A_{ij}}\\ \frac{\delta\phi}{\delta\hat{\varrho}(\mu)} =& \varrho(\mu) = \langle\varrho(\mu|\mathbf{A})\rangle_{p(\mathbf{A})}\end{aligned} \tag{6.2}$$

In the same way as in the previous chapter we change the sum over graphs in the generating function to an expected value over the ER ensemble with the same average degree, $c = \frac{1}{N}\sum_{i=1}^{N} k_i$.

$$\phi[\hat{\varrho}] = \frac{1}{N}\log\left\langle \mathrm{e}^{N\int \mathrm{d}\mu\,\hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})}\prod_{i=1}^{N}\delta_{k_i,\sum_j A_{ij}}\right\rangle_{p_{ER}(\mathbf{A})} + \text{constant}. \tag{6.3}$$

To calculate the generating function we use the same discretization of the integral and replica trick described in the previous chapter. All definitions from the previous chapter follow directly, see equations (5.9)-(5.15). The main difference being the non-uniformity of the degrees, this leads to the following integral,

$$\phi[\hat{\varrho}] = \lim\frac{1}{N}\log\int\prod_{i=1}^{N}\left[\mathrm{d}\boldsymbol{\varphi}^i\nu(\boldsymbol{\varphi}^i)\frac{\mathrm{d}\omega_i}{2\pi}\mathrm{e}^{\mathrm{i}\omega_i k_i}\right]\mathrm{e}^{\sum_{i<j}\log\left[1+\frac{c}{N}\left(\mathrm{e}^{-\mathrm{i}\left[\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}^j+\omega_i+\omega_j\right]}-1\right)\right]}. \tag{6.4}$$

With the same manipulations as in the previous chapter we get to

$$\phi[\hat{\varrho}] = \lim\frac{1}{N}\log\int\mathcal{D}P\mathcal{D}\hat{P}\ \mathrm{e}^{N\mathcal{S}[P,\hat{P}]} + \mathcal{O}(\frac{1}{N^2}), \tag{6.5}$$

$$\begin{aligned}S[P,\hat{P}] =& -\frac{c}{2}(1+\frac{c-2}{2N}) + \mathrm{i}\int\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\omega\ P(\boldsymbol{\varphi},\omega)\hat{P}(\boldsymbol{\varphi},\omega)\\ &+\frac{c}{2}(1+\frac{c}{N})\int\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\omega\mathrm{d}\omega'\ P(\boldsymbol{\varphi},\omega)P(\boldsymbol{\varphi}',\omega')\mathrm{e}^{-\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'-\mathrm{i}\omega-\mathrm{i}\omega'}\\ &-\frac{c^2}{4N}\int\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\omega\mathrm{d}\omega'\ P(\boldsymbol{\varphi},\omega)P(\boldsymbol{\varphi}',\omega')\mathrm{e}^{-2\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'-2\mathrm{i}\omega-2\mathrm{i}\omega'}\\ &-\frac{c}{2N}\int\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\omega\ P(\boldsymbol{\varphi},\omega)\mathrm{e}^{-\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}-2\mathrm{i}\omega} + \sum_k p(k)\log\int\frac{\mathrm{d}\omega\mathrm{d}\boldsymbol{\varphi}}{2\pi}\nu(\boldsymbol{\varphi})\mathrm{e}^{\mathrm{i}\omega k-\mathrm{i}\hat{P}(\boldsymbol{\varphi},\omega)}.\end{aligned} \tag{6.6}$$

The main difference now comes in the appearance of the degree distribution $p(k)$, which assumes that N is large enough for the empirical degree sequence $\frac{1}{N}\sum_i \delta_{k,k_i}$ to be well approximated by $p(k)$. The path forward with the calculation is the same, one needs to introduce the Fourier components of $P(\boldsymbol{\varphi},\omega)$ and $\hat{P}(\boldsymbol{\varphi},\omega)$ w.r.t. ω, and integrate most of

them out. In Appendix D we prove that it can be converted into the following integral over functions $W(\boldsymbol{\varphi})$,

$$\begin{aligned}
\phi[\hat{\varrho}] =& \lim \frac{1}{N} \log \sqrt{\det\left(c\mathbf{U}_1\right)} \int \mathcal{D}W \, \mathrm{e}^{NS_0[W]+S_1[W]} \\
S_0[W] =& \sum_k p(k) \log \int \mathrm{d}\boldsymbol{\varphi} \, \nu(\boldsymbol{\varphi}) \frac{1}{k!} \left(\int \mathrm{d}\boldsymbol{\varphi}' U_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}') W(\boldsymbol{\varphi}') \right)^k \\
& - \frac{c}{2} \int \mathrm{d}\boldsymbol{\varphi} \mathrm{d}\boldsymbol{\varphi}' \, W(\boldsymbol{\varphi}) U_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}') W(\boldsymbol{\varphi}') \\
S_1[W] =& \frac{c^2}{2} \int \mathrm{d}\boldsymbol{\varphi} \mathrm{d}\boldsymbol{\varphi}' \, W(\boldsymbol{\varphi}) U_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}') W(\boldsymbol{\varphi}') - \frac{1}{4} \int \mathrm{d}\boldsymbol{\varphi} \mathrm{d}\boldsymbol{\varphi}' \, r(\boldsymbol{\varphi}) U_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}') r(\boldsymbol{\varphi}') \\
& - \frac{1}{2} \int \mathrm{d}\boldsymbol{\varphi} \, r(\boldsymbol{\varphi}) U_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}).
\end{aligned} \tag{6.7}$$

Where we define the following functions,

$$\begin{aligned}
U_\ell(\boldsymbol{\varphi}, \boldsymbol{\psi}) =& \mathrm{e}^{-\mathrm{i}\ell\boldsymbol{\varphi}\cdot\boldsymbol{\psi}} \\
r(\boldsymbol{\varphi}) =& \sum_k p(k) \frac{k}{c} (k-1) \frac{\nu(\boldsymbol{\varphi}) [\int \mathrm{d}\boldsymbol{\varphi}' U_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}') W(\boldsymbol{\varphi}')]^{k-2}}{\int \mathrm{d}\boldsymbol{\varphi}'' \, \nu(\boldsymbol{\varphi}'') [\int \mathrm{d}\boldsymbol{\varphi}' \, U_1(\boldsymbol{\varphi}'', \boldsymbol{\varphi}') W(\boldsymbol{\varphi}')]^k}
\end{aligned} \tag{6.8}$$

To perform the integral we approximate it using the saddle point approximation, including the Gaussian integral to keep the $\mathcal{O}(1)$ terms as in the previous chapter,

$$\begin{aligned}
\phi[\hat{\varrho}] &= \lim \frac{1}{N} \log \sqrt{\det(c\mathbf{U}_1)} \int \mathcal{D}W \mathrm{e}^{NS_0[W]+S_1[W]} \\
&\sim \lim S_0[W_0] + \frac{1}{N} \lim S_1[W_0] + \frac{1}{N} \lim \log \sqrt{\frac{\det(c\mathbf{U}_1)}{\det(-\boldsymbol{\Gamma})}}.
\end{aligned} \tag{6.9}$$

Where $W_0(\boldsymbol{\varphi})$ satisfies the saddle point equation

$$\frac{\delta S_0}{\delta W}[W_0] = 0. \tag{6.10}$$

This leads to the following equation,

$$\begin{aligned}
W_0(\boldsymbol{\varphi}) =& \sum_k p(k) \frac{k}{c} \frac{F_{k-1}(\boldsymbol{\varphi})}{Z_k} \\
F_k(\boldsymbol{\varphi}) =& \nu(\boldsymbol{\varphi}) \left[\int \mathrm{d}\boldsymbol{\varphi}' U_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}') W_0(\boldsymbol{\varphi}') \right]^k.
\end{aligned} \tag{6.11}$$

The two subleading terms in (6.9) correspond to the subleading terms of the exponent and to the logarithm of the determinant of the Hessian operator of S_0 evaluated at W_0,

denoted by $\mathbf{\Gamma}$. That is,

$$\Gamma(\boldsymbol{\varphi}, \boldsymbol{\psi}) = \left. \frac{\delta^2 S}{\delta W(\boldsymbol{\varphi}) \delta W(\boldsymbol{\psi})} \right|_{W=W_0}. \tag{6.12}$$

As in the previous case it can actually be written down in terms of auxiliary operators of the following three auxiliary operators, $\mathbf{T}, \mathbf{M}, \mathbf{B}$.

$$\begin{aligned}
\mathbf{\Gamma} =& -c\mathbf{U}_1 + c\mathbf{U}_1\mathbf{T} \\
\mathbf{T} =& \mathbf{M} - \mathbf{B} \\
M(\boldsymbol{\varphi}, \boldsymbol{\psi}) =& r_s(\boldsymbol{\varphi}) U_1(\boldsymbol{\varphi}, \boldsymbol{\psi}) \\
B(\boldsymbol{\varphi}, \boldsymbol{\psi}) =& \sum_k p(k) \frac{k^2}{c} \frac{1}{{Z_k}^2} F_{k-1}(\boldsymbol{\varphi}) \cdot \int \mathrm{d}\boldsymbol{\varphi}' U_1(\boldsymbol{\psi}, \boldsymbol{\varphi}') F_{k-1}(\boldsymbol{\varphi}') \\
r_s(\boldsymbol{\varphi}) =& \sum_k p(k) \frac{k}{c} (k-1) \frac{F_{k-2}(\boldsymbol{\varphi})}{Z_k}.
\end{aligned} \tag{6.13}$$

The saddle point equation actually implies $\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}' W_0(\boldsymbol{\varphi}) U_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}') W_0(\boldsymbol{\varphi}') = 1$, so we can rewrite the final result in the following way (leaving out constants).

$$\begin{aligned}
\phi[\hat{\varrho}] =& \lim \Bigg[\sum_k p(k) \log Z_k - \frac{1}{2N} \mathrm{Tr}(\mathbf{B}) - \frac{1}{2N} \mathrm{Tr}(\mathbf{MB}) \\
&+ \frac{1}{4N} \mathrm{Tr}(\mathbf{B}^2) + \frac{1}{N} \sum_{\ell=3}^{\infty} \frac{1}{2\ell} \mathrm{Tr}((\mathbf{M} - \mathbf{B})^\ell) \Bigg]
\end{aligned} \tag{6.14}$$

The series appears after using the log expansion $-\log(1-x) = \sum_{\ell=1}^{\infty} \frac{1}{\ell} x^\ell$ in conjunction with the identity $\log\det(\mathbf{A}) = \mathrm{Tr}\log(\mathbf{A})$. It turns out the $S_1[W_0]$ terms cancel with some of the first terms of the log series. Contrary to the regular degree case, the operators $\mathbf{M}$ and $\mathbf{B}$ do not commute, $\mathbf{MB} \neq \mathbf{BM}$. Therefore no further simplification can be done at this moment before introducing replica symmetry.

6.1.1 Replica symmetric ansatz

We introduce the same replica symmetric (RS) ansatz as before, a superposition of Gaussian distributions with complex diagonal covariance matrix $-\mathrm{i}\mathbf{X}^{-1}$.

RS ansatz:

$$
\begin{aligned}
W_0(\boldsymbol{\varphi}) =& \mathcal{C} \int \mathrm{d}\mathbf{X}\, W(\mathbf{X}) \frac{1}{Z(\mathbf{X})} \mathrm{e}^{-\frac{\mathrm{i}}{2}\boldsymbol{\varphi}\cdot\mathbf{X}\boldsymbol{\varphi}} \\
Z(\mathbf{X}) =& \int \mathrm{d}\boldsymbol{\varphi}\, \mathrm{e}^{-\frac{\mathrm{i}}{2}\boldsymbol{\varphi}\cdot\mathbf{X}\boldsymbol{\varphi}} = \prod_{\mu} \left(\frac{2\pi}{\mathrm{i}x(\mu)}\right)^{\frac{n_\mu}{2}} \left(\overline{\frac{2\pi}{\mathrm{i}x(\mu)}}\right)^{\frac{m_\mu}{2}},
\end{aligned}
\tag{6.15}
$$

where we now have to solve for $W(\mathbf{X})$, this is a distribution of diagonal matrices. The matrices have the same structure as in the previous chapter,

$$
\mathbf{X} = \begin{pmatrix}
x(\mu_1)\mathbf{I}_{n_{\mu_1}} & \mathbf{0} & & \dots & & \mathbf{0} \\
\mathbf{0} & \ddots & \mathbf{0} & & & \\
 & \mathbf{0} & x(\mu_M)\mathbf{I}_{n_{\mu_M}} & \mathbf{0} & & \vdots \\
\vdots & & \mathbf{0} & \overline{x(\mu_1)}\mathbf{I}_{m_{\mu_1}} & \mathbf{0} & \\
 & & & \mathbf{0} & \ddots & \mathbf{0} \\
\mathbf{0} & & \dots & & \mathbf{0} & \overline{x(\mu_M)}\mathbf{I}_{m_{\mu_M}}
\end{pmatrix}.
\tag{6.16}
$$

Therefore the distribution $W(\mathbf{X})$ is over the elements of the diagonal, $\{x(\mu_i)\}_{i=1,\dots,M}$. The distributional equation for $W(\mathbf{X})$ then becomes,

$$
\begin{aligned}
W(\mathbf{X}) =& \frac{1}{\mathcal{C}^2} \sum_k p(k) \frac{k}{c} \frac{Z(\mathbf{X}) f_{k-1}(\mathbf{X})}{\mathcal{Z}_k} \\
\mathcal{C}^2 =& \sum_k p(k) \frac{k}{c} \frac{\mathcal{Z}_{k-1}}{\mathcal{Z}_k} \\
f_k(\mathbf{X}) =& \int \prod_{\ell=1}^{k} \mathrm{d}\mathbf{X}_\ell W(\mathbf{X}_\ell) \cdot \delta(\mathbf{X} + \boldsymbol{\mu} + \sum_{\ell=1}^{k} \mathbf{X}_\ell^{-1}) \\
\mathcal{Z}_k =& \int \mathrm{d}\mathbf{X}\, f_k(\mathbf{X}) Z(\mathbf{X}) = \int \prod_{\ell=1}^{k} \mathrm{d}\mathbf{X}_\ell W(\mathbf{X}_\ell) \cdot Z(-\boldsymbol{\mu} - \sum_{\ell=1}^{k} \mathbf{X}_\ell^{-1}).
\end{aligned}
\tag{6.17}
$$

Note that the matrix δ function can be seen simply as a product over deltas for the diagonal entries. Since each diagonal entry is a complex number, the deltas are acting over the real and imaginary parts. In order for (6.15) to be well defined the complex numbers $x(\mu)$ should have a positive imaginary part, that is $x(\mu) \in \mathbb{H}_{(-)}$ where $\mathbb{H}_{(-)} = \{z \in \mathbb{C} | \mathrm{Im}\, z < 0\}$. Therefore the limit $\varepsilon \to 0$ should only be taken at the very end, to be sure that our

expressions are well defined.

$$\begin{aligned} W(\mathbf{X}) =& W(\{\mathrm{Re}\, x(\mu), \mathrm{Im}\, x(\mu)\}_\mu) \\ \delta(\mathbf{X} - \boldsymbol{\mu} + \sum_{\ell=1}^{k} \mathbf{X}_\ell^{-1}) =& \prod_\mu \delta\left(\mathrm{Re}\left[x(\mu) + \mu_\varepsilon + \sum_{\ell=1}^{k} \frac{1}{x_\ell(\mu)}\right]\right) \\ & \times \delta\left(\mathrm{Im}\left[x(\mu) + \mu_\varepsilon + \sum_{\ell=1}^{k} \frac{1}{x_\ell(\mu)}\right]\right). \end{aligned} \tag{6.18}$$

It is also convenient to introduce the following effective degree distribution, we can see it is properly normalized according to (6.17),

$$\tilde{p}(k) = \frac{1}{\mathcal{C}^2} p(k) \frac{k}{c} \frac{\mathcal{Z}_{k-1}}{\mathcal{Z}_k}. \tag{6.19}$$

With this definition we can rewrite the equation in the following compact way,

$$W(\mathbf{X}) = \sum_k \tilde{p}(k) \frac{1}{\mathcal{Z}_{k-1}} f_{k-1}(\mathbf{X}) Z(\mathbf{X}). \tag{6.20}$$

Using the fact that $Z_k = \mathcal{C}^k \mathcal{Z}_k$ we can rewrite the generating function under the RS ansatz in the following way.

$$\begin{aligned} \phi[\hat{\varrho}] =& \lim \left[-\frac{c}{2} \log \int \mathrm{d}\mathbf{X} \mathrm{d}\mathbf{X}' W(\mathbf{X}) W(\mathbf{X}') \mathcal{K}(\mathbf{X}, \mathbf{X}') + \sum_k p(k) \log \int \mathrm{d}\mathbf{X} f_k(\mathbf{X}) Z(\mathbf{X}) \right. \\ & \left. - \frac{1}{2N} \mathrm{Tr}(\mathbf{B}) - \frac{1}{2N} \mathrm{Tr}(\mathbf{MB}) + \frac{1}{4N} \mathrm{Tr}(\mathbf{B}^2) + \frac{1}{N} \sum_{\ell=3}^{\infty} \frac{1}{2\ell} \mathrm{Tr}((\mathbf{M} - \mathbf{B})^\ell) \right]. \end{aligned} \tag{6.21}$$

Where we have introduced the symmetric function $\mathcal{K}(\mathbf{X}, \mathbf{X}')$,

$$\mathcal{K}(\mathbf{X}, \mathbf{X}') = \frac{Z(\mathbf{X} - \mathbf{X}'^{-1})}{Z(\mathbf{X})} = \prod_\mu \left(\frac{1}{1 - \frac{1}{x(\mu)x'(\mu)}} \right)^{\frac{n_\mu}{2}} \left(\frac{1}{1 - \frac{1}{x(\mu)x'(\mu)}} \right)^{\frac{m_\mu}{2}}, \tag{6.22}$$

and used the identity proven in Appendix D.

$$\mathcal{C}^{-2} = \int \mathrm{d}\mathbf{X} \mathrm{d}\mathbf{X}' W(\mathbf{X}) W(\mathbf{X}') \mathcal{K}(\mathbf{X}, \mathbf{X}'). \tag{6.23}$$

In order to calculate the subleading contributions we only need to calculate two types of traces, pure products of $\mathbf{M}$, that is $\mathrm{Tr}(\mathbf{M}^\ell)$, and of products of $\mathbf{M}$ with insertions of $\mathbf{B}$'s,

such as $\mathrm{Tr}(\mathbf{M}^3\mathbf{B}^2\mathbf{M}^5\mathbf{B}^6)$. Their exact expressions can be found in Appendix D. To avoid unnecessary repetition we will only show them in the main text after taking the imaginary replica limit

6.1.2 Imaginary replica limit

We can now easily take the replica limit according to the convention (4.50) discussed in Chapter 4,

$$\begin{aligned}\lim \prod_{\mu} f(\mu)^{n_\mu}\overline{f(\mu)}^{m_\mu} &= \lim_{\Delta\to 0}\lim_{n_\mu\to \mathrm{i}\frac{\Delta}{\pi}\hat{\varrho}'(\mu)}\lim_{m_\mu\to -n_\mu} \mathrm{e}^{\sum_\mu\left[n_\mu\log f(\mu)+m_\mu\log\overline{f(\mu)}\right]}\\ &= \mathrm{e}^{\frac{2}{\pi}\int \mathrm{d}\mu\ \hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log f(\mu)}.\end{aligned} \tag{6.24}$$

In the limit $\Delta\to 0$ ($M\to\infty$) we recover the integral over μ. This means that our theory becomes functional in form. That is we will now look for a distribution over the space of functions $x(\mu)$ with $\mu\in[-\mu_0,\mu_0]$. We will denote this change in the following way.

$$\mathrm{d}\mathbf{X}\to\mathcal{D}\boldsymbol{x}=\prod_\mu \mathrm{d}x(\mu)=\prod_\mu \mathrm{dRe}[x(\mu)]\mathrm{dIm}[x(\mu)] \quad \text{and} \quad W(\mathbf{X})\to W[\boldsymbol{x}]. \tag{6.25}$$

We can then write down the saddle point equation as a distributional equation for the distribution of functions $W[\boldsymbol{x}]$.

$$\begin{aligned}W[\boldsymbol{x}] &= \sum_k \tilde{p}(k)\frac{\mathrm{e}^{\mathcal{A}[\boldsymbol{x}]}f_{k-1}[\boldsymbol{x}]}{\mathcal{Z}_{k-1}}\\ \mathcal{A}[\boldsymbol{x}] &= -\frac{1}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu)\\ f_k[\boldsymbol{x}] &= \int\prod_{\ell=1}^{k}\mathcal{D}\boldsymbol{x}_\ell W[\boldsymbol{x}_\ell]\cdot\prod_\mu\delta\left(x(\mu)+\mu+\sum_{\ell=1}^{k}\frac{1}{x_\ell(\mu)}\right)\\ \mathcal{Z}_k &= \int\mathcal{D}\boldsymbol{x}f_k[\boldsymbol{x}]\mathrm{e}^{\mathcal{A}[\boldsymbol{x}]}\\ \tilde{p}(k) &= \frac{1}{\mathcal{C}^2}p(k)\frac{k}{c}\frac{\mathcal{Z}_{k-1}}{.}\end{aligned} \tag{6.26}$$

Note that these equations are very reminiscent of equations derived for graphs sampled from the configuration model, [39]. Nevertheless they have several key differences:

- The unknown is a distribution over functions, $x(\mu)$, and not over a finite number of complex numbers. Specifically, $W[\boldsymbol{x}]$ is a distribution over curves in the complex lower-half plane, that is functions $x:[-\mu_0,\mu_0]\subseteq\mathbb{R}\to\mathbb{H}_{(-)}$.

- The effective degree distribution is not simply $p(k)k/c$, it is tilted with a factor $\mathcal{Z}_{k-1}/\mathcal{Z}_k$.
- The distribution $f_{k-1}[\boldsymbol{x}]$ is also weighted with an extra factor $\mathrm{e}^{\mathcal{A}[\boldsymbol{x}]}$.

Similar equations with the extra weight fact have also been observed for weighted graphs in [54–56, 58], but without the functional nature. Before discussing possible ways of solving, we will finish writing down all the elements of the theory.

The generating function becomes

$$\begin{aligned}\phi[\hat{\varrho}] =& -\frac{c}{2}\log\int \mathcal{D}\boldsymbol{x}\mathcal{D}\boldsymbol{x}' W[\boldsymbol{x}]W[\boldsymbol{x}']\mathrm{e}^{-\frac{1}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log(1-\frac{1}{x(\mu)x'(\mu)})}\\ &+\sum_k p(k)\log\int \mathcal{D}\boldsymbol{x} f_k[\boldsymbol{x}]\mathrm{e}^{\mathcal{A}[\boldsymbol{x}]}+\tilde{\phi}_0(N)\\ &-\frac{1}{2N}\lim\mathrm{Tr}(\mathbf{B})-\frac{1}{2N}\lim\mathrm{Tr}(\mathbf{MB})+\frac{1}{4N}\lim\mathrm{Tr}(\mathbf{B}^2)+\frac{1}{N}\sum_{\ell=3}^{\infty}\frac{1}{2\ell}\lim\mathrm{Tr}[(\mathbf{M}-\mathbf{B})^\ell].\end{aligned} \tag{6.27}$$

Here we have included all the previously ignored constant (w.r.t $\hat{\varrho}$) terms in a single term $\tilde{\phi}_0(N)$. As we mentioned before, all the subleading terms can be written in terms of just two types of traces. It turns our that they both have a very clear interpretation and pictorial representation, as it was already hinted at the previous chapter. All the traces coming from expanding the order ℓ term, $\mathrm{Tr}[(\mathbf{M}-\mathbf{B})^\ell$, correspond to subgraphs of ℓ edges. The pure traces of $\mathbf{M}$, $\mathrm{Tr}(\mathbf{M}^\ell)$, correspond to loops of length ℓ embedded in a graph, where the degree of each node is drawn with probability distribution $\tilde{p}(k)$, (6.19). The mixed traces with $\mathbf{M}$'s and $\mathbf{B}$'s (including the case of only $\mathbf{B}$'s), $\mathrm{Tr}(\prod_{t=1}^{s}\mathbf{M}^{\ell_t-1}\mathbf{B}^{m_t+1})$, correspond to embedded chains, one for each $\mathbf{B}$.

For the first type we get,

$$\begin{aligned}\lim\mathrm{Tr}(\mathbf{M}^\ell) =& \left[\prod_{p=1}^{\ell}\sum_{k_p}\tilde{p}(k_p)(k_p-1)\right]\exp\left(\ell R(\{k_p\}_{p=1,\dots,\ell})\right)\\ R(\{k_p\}_{p=1,\dots,\ell}) =& \frac{1}{\ell}\log\int\prod_{p=1}^{\ell}\mathcal{D}\boldsymbol{x}_p f_{k_p-2}[\boldsymbol{x}_p]\exp\left(\ell\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}g_\ell(\mu|\{x_p(\mu)\}_{p=1,\dots,\ell})\right)\\ &-\frac{1}{\ell}\sum_{p=1}^{\ell}\log\int\mathcal{D}\boldsymbol{x} f_{k_p-1}[\boldsymbol{x}]\mathrm{e}^{-\frac{1}{\pi}\int\mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu)}\\ g_\ell(\mu|\{x_p(\mu)\}_{p=1,\dots,\ell}) =& \frac{2}{\pi\ell}\mathrm{Im}\log Z(\mu|\mathbf{A}^*_{\ell,\mu}).\end{aligned} \tag{6.28}$$

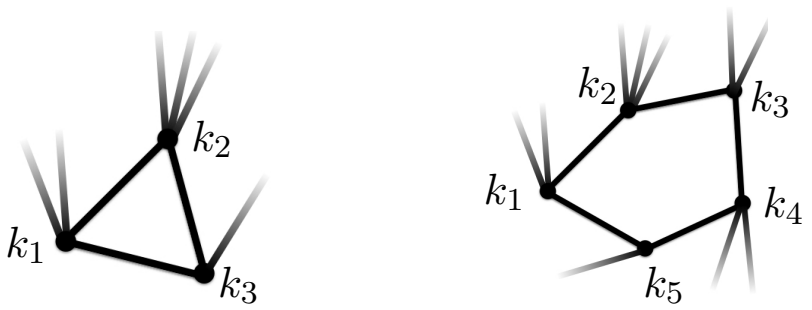


Figure 6.1: Pictorial representation of elements containing only $\mathbf{M}$ in the series in (6.27). Each node of the loop has degree k_p sampled form $\tilde{p}(k_p)$. Each node is connected to two other nodes in the loop and receives $k_p - 2$ fields of value $-1/x(\mu)$ from the rest of the graph, for this reason they are sampled from $W[\boldsymbol{x}]$.

Where we have defined $\mathbf{A}^*_{\ell,\mu}$ to be the family of adjacency matrices of a ring of length ℓ with complex fields $\mu + x_p(\mu)$ on the diagonal.

$$(\mathbf{A}^*_{\ell,\mu})_{p,p'} = \delta_{p,p'-1} + \delta_{p,p'+1} + (\mu + x_p(\mu))\delta_{p,p'} \quad \text{(with } p \text{ mod } \ell\text{)}. \tag{6.29}$$

Note that since the distribution of $\boldsymbol{x}_p$ is distributed according to $f_{k_p-2}[\boldsymbol{x}_p]$ this means that in the diagonal there are $k_p - 2$ fields with values $-(x(\mu))^{-1}$ distributed according to $W[\boldsymbol{x}]$. This has a natural interpretation as an isolated embedded ring receiving "messages" from an infinite tree-like graph with degree distribution $p(k)$. See Figure 6.1. This points towards a possible cavity derivation of this expression.

It is very important to notice that $\frac{\mathrm{d}}{\mathrm{d}\mu} g_\ell(\mu|\{x_p(\mu)\}_{p=1,\dots,\ell})$ *looks* like the EJ formula (1.30) for a finite $\ell \times \ell$ matrix $\mathbf{A}^*_{\ell,\mu}$, but it is *not*. The reason being that in this case we have a *different* matrix $\mathbf{A}^*_{\ell,\mu}$ for each value of μ. This means that $\frac{\mathrm{d}}{\mathrm{d}\mu} g_\ell(\mu)$ is typically a continuous function and not a discrete sum of δ's. In the previous chapter there was only one $g_\ell(\mu)$ and it was shown to be a continuous function, examples of $\frac{\mathrm{d}}{\mathrm{d}\mu} g_\ell(\mu|\{x_p(\mu)\}_{p=1,\dots,\ell})$'s can be seen in Figure 6.4.

For the other type of traces we can also write down the general expression,

$$\mathrm{Tr}(\prod_{t=1}^{s}\mathbf{M}^{\ell_t-1}\mathbf{B}^{m_t+1}) = \left[\prod_{t=1}^{s}\prod_{p_t=1}^{\ell_t+m_t}\sum_{k_{p_t}}\tilde{p}(k_{p_t})\right]\prod_{t=1}^{s}\left[\prod_{p_t=1}^{\ell_t-1}(k_{p_t}-1)\prod_{p_t=\ell_t}^{\ell_t+m_t}k_{p_t}\right] \times \exp\left(\sum_{t=1}^{s}\left[(\ell_t+1)C(k_{\ell_{t-1}+m_{t-1}},\{k_{p_t}\}_{p_t=1,\dots,\ell_t}) + \sum_{p_t=\ell_t+1}^{\ell_t+m_t}2C(k_{p_t-1},k_{p_t})\right]\right). \tag{6.30}$$

Where we have defined,

$$\begin{aligned} C(k,\{k_p\}_{p=1,\dots,\ell}) =& \frac{1}{\ell+1}\log\int\mathcal{D}\boldsymbol{x}f_{k-1}[\boldsymbol{x}]\mathcal{D}\boldsymbol{x}_\ell f_{k_\ell-1}[\boldsymbol{x}_\ell] \\ &\times\prod_{p=1}^{\ell-1}\mathcal{D}\boldsymbol{x}_p f_{k_p-2}[\boldsymbol{x}_p]\exp\left(\ell\int\mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}c_\ell(\mu|x,\{x_p\}_{p=1,\dots,\ell})\right) \\ &-\frac{1}{\ell+1}\log\int\mathcal{D}\boldsymbol{x}f_k[\boldsymbol{x}]\mathrm{e}^{-\frac{1}{\pi}\int\mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu)} \\ &-\frac{1}{\ell+1}\sum_{p=1}^{\ell}\log\int\mathcal{D}\boldsymbol{x}f_{k_p-1}[\boldsymbol{x}]\mathrm{e}^{-\frac{1}{\pi}\int\mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu)}, \\ c_\ell(\mu|x,\{x_p\}_{p=1,\dots,\ell}) =& \frac{2}{\pi(\ell+1)}\mathrm{Im}\log Z(\mu|\boldsymbol{C}^*_{\ell+1,\mu}). \end{aligned} \tag{6.31}$$

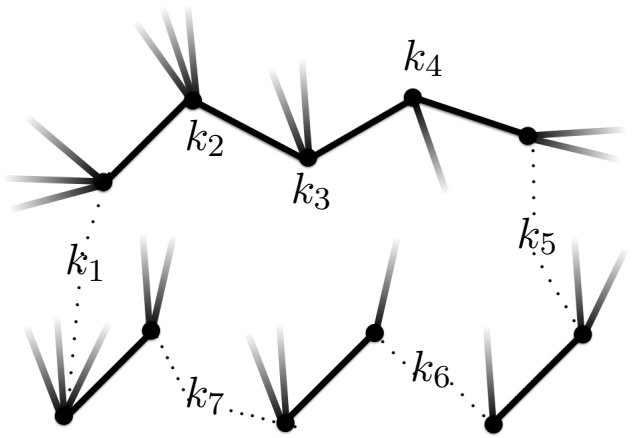
In this case $\boldsymbol{C}^*_{\ell+1,\mu}$ corresponds to an embedded open chain of length ℓ (with $\ell+1$ nodes), with $k-2$ complex fields in the inner nodes, and $k-1$ fields on the first and last nodes.

$$\boldsymbol{C}^*_{\ell+1,\mu} = \delta_{p,p'-1}+\delta_{p,p'+1}+(\mu+x_p(\mu))\delta_{p,p'}. \tag{6.32}$$

Similarly to $g_\ell(\mu)$ of (6.28), $c_\ell(\mu)$ looks like the EJ formula for a single matrix but it is actually a continuous function due to the presence of the complex fields. Examples in Figure 6.4.

Therefore the terms $\mathrm{Tr}(\prod_{t=1}^{s}\mathbf{M}^{\ell_t-1}\mathbf{B}^{m_t+1})$ contribute with s open chains of lengths $\{\ell_t\}_{t=1,\dots,s}$ each, and $\sum_{t=1}^{s}m_t$ links (chains of length 1). See Figure 6.2 for visual representation of this terms. Note that even though they are not connected, the endpoints of the chains share the same degree.

All the terms in the traces can then be represented either loops as in Figure 6.1 or by combinations of diagrams as in Figure 6.2.



$$\mathrm{Tr}(\mathbf{M}^3\mathbf{B}^4)$$

Figure 6.2: Pictorial representation of elements containing $\mathbf{M}$ and $\mathbf{B}$ in the series in (6.27). The diagram contains ℓ edges and can be constructed by building a loop of length ℓ but each $\mathbf{B}$ disconnects a new chain. Each node of the loop has degree k_p sampled form $\tilde{p}(k_p)$. Each internal node is connected to two other nodes in the chain and therefore receives $k_p - 2$ fields of value $-1/x(\mu)$, the end points receive $k_p - 1$. All the fields are sampled from $W[\boldsymbol{x}]$. The dotted lines indicate the nodes that share the same degree value, they do not represent edges.

6.1.3 Formula for the spectrum

Finally we show the formula for the average spectrum, it is obtained by taking the functional derivative of the generating function (6.27) w.r.t $\hat{\varrho}(\mu)$. Given that $W[\boldsymbol{x}]$ is a stationary point of $S_0[W]$ we only need to take the explicit functional derivatives of the leading order terms,

$$\begin{aligned}\varrho(\mu) = \frac{\delta\phi[\hat{\varrho}]}{\delta\hat{\varrho}(\mu)} = &-\frac{c}{2}\int \mathcal{D}\boldsymbol{x}\mathcal{D}\boldsymbol{x}' W_{\mathcal{K}}[\boldsymbol{x},\boldsymbol{x}']\left(-\frac{1}{\pi}\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log(1-1/(x(\mu)x'(\mu)))\right)\\ &+\sum_k p(k)\int \mathcal{D}\boldsymbol{x} W_k[\boldsymbol{x}]\left(-\frac{1}{\pi}\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu)\right)\\ &+\frac{1}{N}\frac{\delta}{\delta\hat{\varrho}(\mu)}\lim\left[-\frac{1}{2}\mathrm{Tr}(\mathbf{B})-\frac{1}{2}\mathrm{Tr}(\mathbf{MB})+\frac{1}{4}\mathrm{Tr}(\mathbf{B}^2)+\sum_{\ell=3}^{\infty}\frac{1}{2\ell}\mathrm{Tr}((\mathbf{M}-\mathbf{B})^\ell)\right].\end{aligned} \tag{6.33}$$

Where we have defined the following distributions.

$$W_{\mathcal{K}}[\boldsymbol{x},\boldsymbol{x}'] = \frac{W[\boldsymbol{x}]W[\boldsymbol{x}']\mathrm{e}^{-\frac{1}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log(1-1/(x(\mu)x'(\mu)))}}{\int \mathcal{D}\boldsymbol{x}\mathcal{D}\boldsymbol{x}'W[\boldsymbol{x}]W[\boldsymbol{x}']\mathrm{e}^{-\frac{1}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log(1-1/(x(\mu)x'(\mu)))}}$$
$$W_k[\boldsymbol{x}] = \frac{f_k[\boldsymbol{x}]\mathrm{e}^{-\frac{1}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu)}}{\int \mathcal{D}\boldsymbol{x}\, f_k[\boldsymbol{x}]\mathrm{e}^{-\frac{1}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu)}}. \tag{6.34}$$

For the subleading terms the story is not so simple, as we cannot in general discard the $\delta W/\delta\hat{\varrho}(\mu)$ terms coming from the chain rule. Although it is possible to write an implicit equation for this derivative starting from (6.26), we will not pursue this path. It will not be necessary as we will show in the following section that $\delta W/\delta\hat{\varrho}(\mu) = 0$ for our case of interest, and we will still be able to extract nontrivial results.

6.2 State of the theory and simple tests

At this point we can sum up the general state of the theory. We have developed a general theory to calculate the generating function $\phi[\hat{\varrho}]$ for the spectrally constrained ensemble (6.1) in leading order in N, the $\mathcal{O}(1)$ contributions, and in subleading order the $\mathcal{O}(1/N)$ contributions. This theory suggest the next approach:

- Choose a function $\hat{\varrho}$ such that $\hat{\varrho}(\mu) = \mathcal{O}(1)$, such that it represents the desired observable. For example, for a bias like $\alpha\mathrm{Tr}(\mathbf{A}^\ell)$ choose $\hat{\varrho}(\mu) = \alpha\mu^\ell$.
- Solve the saddle point equation (6.26) for this choice of $\hat{\varrho}$.
- Evaluate *both* $\mathcal{O}(1)$ and $\mathcal{O}(1/N)$ terms of the generating function (6.27).
- Evaluate the leading order term of the spectral density according to (6.33).
- *If* it is possible to calculate $\delta W/\delta\hat{\varrho}(\mu)$ then it is also possible to calculate the $\mathcal{O}(1/N)$ corrections to the spectral density.

It is *very important* to point out that this theory *does not* guarantee that the most important subleading terms will be $\mathcal{O}(1/N)$. It could be the case that other contributions of $\mathcal{O}\left(1/N^\delta\right)$ with $\delta < 1$ exist. They would not correspond to finite size corrections due to loops or chains as the ones described so far. While we have not observed these anomalous finite size corrections for any of the models presented in this thesis, it is a possibility that

should not be discarded. They have been observed in low temperature spin glass models on random regular graphs in [72].

The only rigorous proof known to the author is for regular graphs, $p(k) = \delta_{k,q}$, and for the $\hat{\varrho}(\mu) = 0$ case. In [66] it is proved that fluctuations are $\mathcal{O}(1/N)$ and that they correspond exactly to the ones predicted by this theory, as shown in Chapter 5.

6.2.1 Solving the equations numerically

While finding a general solution of (6.26) is most likely an impossible task, it is possible to write down an algorithm to numerically approximate its solutions with a variation of the population dynamics algorithm, [39]. The main difference will come from the introduction of weights due to the presence of the imaginary replicas. We have already introduced the effective degree distribution $\tilde{p}(k)$ in (6.19), we will see that the other part can also be interpreted as a probability distribution, the conditional distribution of $\boldsymbol{x}$ given $k > 0$.

$$P(\boldsymbol{x}|k) = \frac{f_{k-1}[\boldsymbol{x}]\mathrm{e}^{\mathcal{A}[\boldsymbol{x}]}}{\int \mathcal{D}\boldsymbol{x}' f_{k-1}[\boldsymbol{x}']\mathrm{e}^{\mathcal{A}[\boldsymbol{x}']}} = \frac{1}{\mathcal{Z}_{k-1}} f_{k-1}[\boldsymbol{x}]\mathrm{e}^{\mathcal{A}[\boldsymbol{x}]}. \tag{6.35}$$

Note $P(\boldsymbol{x}|1) = \delta(\boldsymbol{x} + \boldsymbol{\mu})$, since $\mathcal{A}[-\boldsymbol{\mu}] = \hat{\varrho}(0)$.

With this we can then define a dynamical process for a population of function $x(\mu)$ such that it should converge to a population distributed according to $W[\boldsymbol{x}]$ in (6.26).

The weighted population dynamics algorithm for the functional equation (6.26) is the following one:

1. Choose a finite interval $[-\mu_0, \mu_0]$ and divide in to $M-1$ equal intervals of length $\Delta = 2\mu_0/(M-1)$. Denote by $\{\mu_i\}_{i=1,\dots,M}$ the edges of such intervals. Also choose a small regularization parameter $0 < \varepsilon << 1$ and a maximum degree q_0.

2. Create a random population of L complex vectors with positive imaginary part of complex dimension M. That is $\{\vec{\boldsymbol{x}}_a\}_{a=1,\dots,L}$ with $\vec{\boldsymbol{x}}_a \in \mathbb{H}_{(-)}^M$, where $\mathbb{H}_{(-)} = \{z \in \mathbb{C} | \mathrm{Im} z < 0\}$. They constitute the discrete approximation of a sample from $W[\boldsymbol{x}]$.

   Averages are calculated empirically over the population,

$$\int \prod_{p=1}^{\ell} \mathcal{D}\boldsymbol{x}_p W[\boldsymbol{x}_p] f[\{\boldsymbol{x}_p\}] \approx \prod_{p=1}^{\ell} \sum_{a_p} L^{-\ell} f(\{\vec{\boldsymbol{x}}_{a_p}\}_{p=1,\dots,\ell}), \tag{6.36}$$

and integrals are approximated in the following way,

$$\int_{-\mu_0}^{\mu_0} \mathrm{d}\mu f(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}g(x(\mu)) \approx f(\mu_0)g(x(\mu_0)) - f(-\mu_0)g(x(-\mu_0)) - \sum_{i=1}^{M-1}\frac{g(x(\mu_i))+g(x(\mu_{i+1}))}{2}(f(\mu_{i+1})-f(\mu_i)). \quad (6.37)$$

Integrals are calculated in this way inspired by the integration by parts theorem for Stieltjes integrals, [73]. In this way we avoid taking a numerical derivative of the noisy object $g(x(\mu))$ and take them instead of the known function $f(\mu)$.

3. Calculate $\mathcal{Z}_{k-1}$ and $\mathcal{Z}_k$ for k such that $p(k)k > 0$ and $k \leq q_0$ by averaging empirically $\mathrm{e}^{\mathcal{A}[\boldsymbol{x}]}$.

4. Draw a degree k according to $\tilde{p}(k) = p(k)\frac{k}{c}\frac{\mathcal{Z}_{k-1}}{\mathcal{Z}_k}$.

5. Choose $k-1$ elements at random from $\{\vec{\boldsymbol{x}}_a\}_{a=1,\dots,L}$ and calculate a new $\vec{\boldsymbol{x}}$ according to $x(\mu) = -\mu_\varepsilon + \sum_{\ell=1}^{k-1} 1/x_\ell(\mu)$.

6. Calculate $w = \mathrm{e}^{\mathcal{A}[\boldsymbol{x}]}/\mathcal{Z}_{k-1}$. Replace $\lfloor w \rfloor$ elements from $\{\vec{\boldsymbol{x}}_a\}_{a=1,\dots,L}$ with $\vec{\boldsymbol{x}}$. Replace an additional random element with probability $w - \lfloor w \rfloor$. Where $\lfloor a \rfloor$ denotes the floor function.

7. Repeat steps 3. to 6. T times until equilibration.

Once one has a large population of $\{\vec{\boldsymbol{x}}_a\}_{a=1,\dots,L}$, any observable can be calculated with the same prescription as in point 2.

This algorithm is an adaptation of the one presented in [55] for a similar family of random matrices. In that case the functional formalism was not used but it also contained imaginary replicas. For this reason similar equations with weights were obtained, therefore the authors of [55] proposed the weighted population dynamics algorithm.

We include the algorithm for completeness, but as a matter of fact for the purpose of this work it will not be necessary to use it as it turns out the trivial solutions (those with $\hat{\varrho} = 0$) are enough for the cases studied here. This was shown analytically for regular graphs in Chapter 5.

6.2.2 Test of the leading order with $\hat{\varrho}(\mu) = 0$

As a first test for the theory we will recover the theory for the spectral density of random graphs with an specified degree distribution sample uniformly, [39]. This is achieved by

simply setting $\hat{\varrho}(\mu) = 0$ in (6.1). This turns the saddle point equation (6.26) into

$$W[\boldsymbol{x}] = \sum_k p(k) \frac{k}{c} f_{k-1}[\boldsymbol{x}]. \tag{6.38}$$

In order to link the formula for the spectrum presented here (6.33) with the one of [39] we need to introduce the marginal distributions,

$$w(x|\mu) = \int \prod_{\mu' \neq \mu} dx(\mu')\, W[\boldsymbol{x}]. \tag{6.39}$$

The marginals satisfy the same equation,

$$w(x|\mu) = \sum_k p(k) \frac{k}{c} \int \prod_{\ell=1}^{k-1} \mathrm{d}x_\ell w(x_\ell|\mu) \delta \left[x + \mu + \sum_{\ell=1}^{k-1} \frac{1}{x_\ell} \right]. \tag{6.40}$$

This is exactly the equation derived in [39], with just minor convention differences.

Now we need to show the formula for the spectrum is also the same as the one in [39]. Note that when $\hat{\varrho}(\mu) = 0$ we get the following simplifications due to normalization of $W[\boldsymbol{x}]$,

$$\begin{aligned} W_{\mathcal{K}}[\boldsymbol{x}, \boldsymbol{x}'] =& W[\boldsymbol{x}] W[\boldsymbol{x}'], \\ W_k[\boldsymbol{x}] =& f_k[\boldsymbol{x}]. \end{aligned} \tag{6.41}$$

We will only focus on the leading order term of (6.33) for $\hat{\varrho}(\mu) = 0$. We will denote it as

$$\varrho_0(\mu) = \lim_{N \to \infty} \frac{\delta \phi}{\delta \hat{\varrho}(\mu)}[0] \tag{6.42}$$

If we take out of the integrals the μ derivative in (6.33) we can write everything in terms of integrals over marginals (6.39),

$$\varrho_0(\mu) = \frac{1}{\pi} \frac{\mathrm{d}}{\mathrm{d}\mu} \left[\frac{c}{2} \int \mathrm{d}x \mathrm{d}x' w(x|\mu) w(x'|\mu) \mathrm{Im} \log(1 - \frac{1}{xx'}) - \sum_k p(k) \int \mathrm{d}x f_k(x|\mu) \mathrm{Im} \log x \right]. \tag{6.43}$$

To proceed we need to assume that $\frac{\partial}{\partial \mu} w(x|\mu) = \partial_\mu w(x|\mu)$ exists. Then, we can carry out

all of the derivatives with respect to μ on the RHS, giving

$$\begin{aligned}
\varrho_0(\mu) =& -\frac{1}{\pi}\mathrm{Im}\left[-c\int \mathrm{d}x\mathrm{d}x' w(x|\mu)\partial_\mu w(x'|\mu)\log(1-\frac{1}{xx'})\right.\\
&\left.+\sum_k p(k)\frac{\mathrm{d}}{\mathrm{d}\mu}\int\prod_{\ell=1}^{k}\mathrm{d}x_\ell w(x_\ell|\mu)\log(-\mu-\sum_{\ell=1}^{k}\frac{1}{x_\ell})\right]\\
=& -\frac{1}{\pi}\mathrm{Im}\left[-c\int \mathrm{d}x\mathrm{d}x' w(x|\mu)\partial_\mu w(x'|\mu)\log(1-\frac{1}{xx'})\right.\\
&+\sum_k p(k)k\int \mathrm{d}x'\partial_\mu w(x'|\mu)\prod_{\ell=1}^{k-1}\mathrm{d}x_\ell w(x_\ell|\mu)\log(-\mu-\sum_{\ell=1}^{k-1}-\frac{1}{x_\ell}-\frac{1}{x'})\\
&\left.-\sum_k p(k)\int\prod_{\ell=1}^{k}\mathrm{d}x_\ell w(x_\ell|\mu)\frac{1}{-\mu-\sum_{\ell=1}^{k}\frac{1}{x_\ell}}\right]\\
=& -\frac{1}{\pi}\mathrm{Im}\left[-c\int \mathrm{d}x\mathrm{d}x' w(x|\mu)\partial_\mu w(x'|\mu)\log(1-\frac{1}{xx'})\right.\\
&+\sum_k p(k)k\int \mathrm{d}x'\partial_\mu w(x'|\mu)\int \mathrm{d}x f_{k-1}(x|\mu)\log(x-\frac{1}{x'})\\
&\left.-\sum_k p(k)\int \mathrm{d}x f_k(x|\mu)\frac{1}{x}\right]\\
=& -\frac{1}{\pi}\mathrm{Im}\left[c\int \mathrm{d}x w(x|\mu)\log x\int \mathrm{d}x'\partial_\mu w(x'|\mu)-\sum_k p(k)\int \mathrm{d}x f_k(x|\mu)\frac{1}{x}\right]\\
=&\frac{1}{\pi}\sum_k p(k)\int \mathrm{d}x f_k(x|\mu)\mathrm{Im}\frac{1}{x}.
\end{aligned}\tag{6.44}$$

Here we have used the saddle point equation and the fact that $\int \mathrm{d}x\partial_\mu w(x|\mu) = 0$ due to normalization of $w(x|\mu)$. This again is exactly the expression for the spectral density shown in [39].

In particular this implies that we have found an analytic expression not only for the asymptotic spectral density, but also for the asymptotic cumulative density function (CDF).

$$\begin{aligned}
R_0(\mu) =& \int_{-\infty}^{\mu}\mathrm{d}s\,\varrho_0(s)\\
=&\frac{1}{\pi}\left[\frac{c}{2}\int \mathrm{d}x\mathrm{d}x' w(x|\mu)w(x'|\mu)\mathrm{Im}\log(1-\frac{1}{xx'})-\sum_k p(k)\int \mathrm{d}x f_k(x|\mu)\mathrm{Im}\log x\right].
\end{aligned}\tag{6.45}$$

This result is analogous to the one found in [54] for sparse graphs as well.

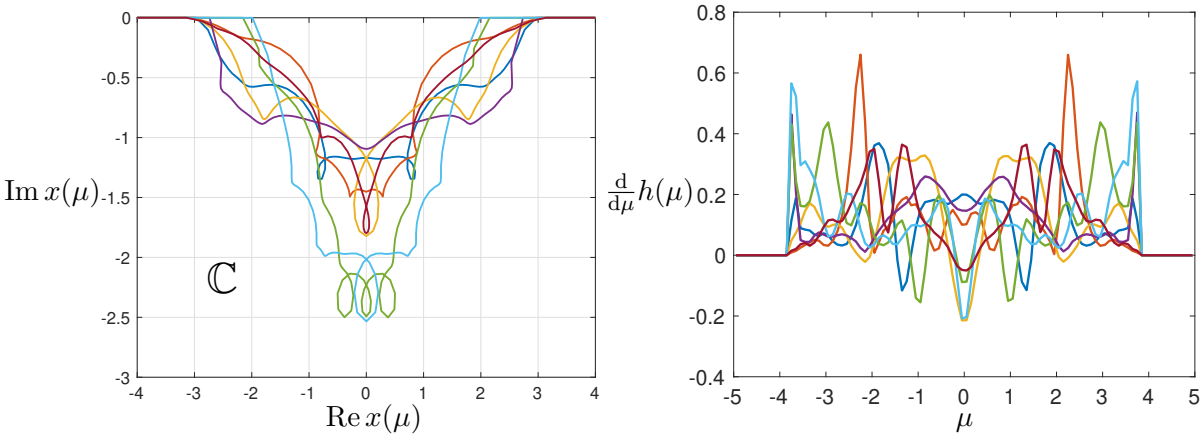


Figure 6.3: Samples of $W[\boldsymbol{x}|0]$ generated with the population dynamics algorithm described in the text. On the left panel we show example of curves $x(\mu) \in \mathbb{C}$, notice these are not plots but they are actually the curves themselves on the complex plane. On the right panel we plot the corresponding spectral contributions, $\frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu) = -\frac{1}{\pi}\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu)$, of each sample with the same color.

6.3 Applying the general theory for $\hat{\varrho}(\mu) = \alpha\mu^3$

We now discuss the case of a triangle bias, which in spectral terms means $\hat{\varrho}(\mu) = \alpha\mu^3$. In previous chapters we have studied this model at length through MCMC sampling methods and theoretically with combinatorial arguments. Our aim is to use the developed replica theory to recover the same $\mathcal{O}(1/N)$ formula for $m(\alpha)$ as in Chapter 3 and to additionally provide a formula for the average spectral density on that regime. This implies a generalization of the previous chapter.

6.3.1 The generating function for $\hat{\varrho}(\mu) = \alpha\mu^3$

As we showed in Chapters 2 and 3, we do not expect a bias of this kind to have any effect on the leading order of this model. This implies that it could be that solving our functional saddle point equation with $\hat{\varrho}(\mu) = \alpha\mu^3$ gave back the same solutions as if $\hat{\varrho}(\mu) = 0$. Actually looking at the structure of the saddle point equation (6.26), it is not hard to convince oneself that solutions to the unbiased equation (6.38) should also be solutions of (6.26). Given that it is well known that graphs sampled from the configuration model have a symmetric spectral density asymptotically (because they have only a finite number of loops, [5]), it is just natural to imagine the contributions of each functional message should be symmetric as well. This can be seen in Figure 6.3 where we show typical samples from $W[\boldsymbol{x}]$. These are curves in the complex plane. Introducing the notation

$h(\mu) = -\frac{1}{\pi}\mathrm{Im}\log x(\mu)$, we also show $\frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu)$ in Figure 6.3. We can see that as expected they are symmetric in all cases. If we introduce the notation $W[\boldsymbol{x}|\hat{\varrho}]$ to mean the solution of (6.26) for functional bias $\hat{\varrho}$, we claim that

$$-\frac{1}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu) = 0 \qquad \begin{array}{l}\text{if } x(\mu) \text{ is a sample of } W[\boldsymbol{x}|0] \\ \text{and } \hat{\varrho}(\mu) \text{ and odd function.}\end{array} \tag{6.46}$$

We have tested this claim numerically with the population dynamics algorithm described in 6.2.1. For simplicity we chose the simplest nontrivial degree distribution, $p(k) = \frac{1}{2}(\delta_{k,3}+\delta_{k,5})$. For a population of $L = 2\times 10^4$ vectors, $\varepsilon = 10^{-13}$, and $\Delta = 0.1$, we find that the empirical average of (6.46) is $\frac{1}{L}\sum_{a=1}^{L} -\frac{1}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x_a(\mu) = 0.001(2)$. Note that (6.46) is a claim *for all* $x(\mu)$ sampled from $W[\boldsymbol{x}]$, the small standard deviation of the empirical average is consistent with this fact. This also gives us an idea that our theory with these parameters for the population dynamics has approximately that accuracy.

Knowing (6.46), we are then sure that the solution for $\hat{\varrho}(\mu) = 0$ will also satisfy the equation for $\hat{\varrho}(\mu) = \alpha\mu^3$. In our notation,

$$W[\boldsymbol{x}|0] = W[\boldsymbol{x}|\alpha\mu^3] \qquad \text{for all } \alpha \tag{6.47}$$

Note that we have already observed this for regular graphs in the previous chapter. In that case it was easy to prove with the exact solution which was simply a δ-function, here we claim that it actually extends to other degree distributions $p(k)$. As mentioned in Chapter 3, we expect this to be true at least for bounded degree distributions. For unbounded degree distributions it is still not clear if there may perhaps be loops asymptotically for some values of $\alpha = \mathcal{O}(1)$, as discussed in (3). Nevertheless even if that were the case, it is unlikely that these set of equations would describe the situation as we would not expect an equation with such a strong resemblance to the cavity method to describe it. As a simple indication, the equations found in this chapter (6.28) required the eigenvalues of actual loops to give loopy contributions.

The amazing thing is that even though the identity (6.47) implies that the leading order of $\phi[\hat{\varrho}(\mu) = \alpha\mu^3]$ is the same as the $\hat{\varrho}(\mu) = 0$ case, the $\mathcal{O}(1/N)$ can be modified by the presence of this bias, similarly to the regular case explained in the previous chapter.

Given the interpretation of $g_\ell(\mu|\{x_p\}_{p=1,\ldots,\ell})$ as the spectral contribution of a loop of length ℓ embedded in the graph, we expect the moments of $g_\ell(\mu)$ to reflect the average

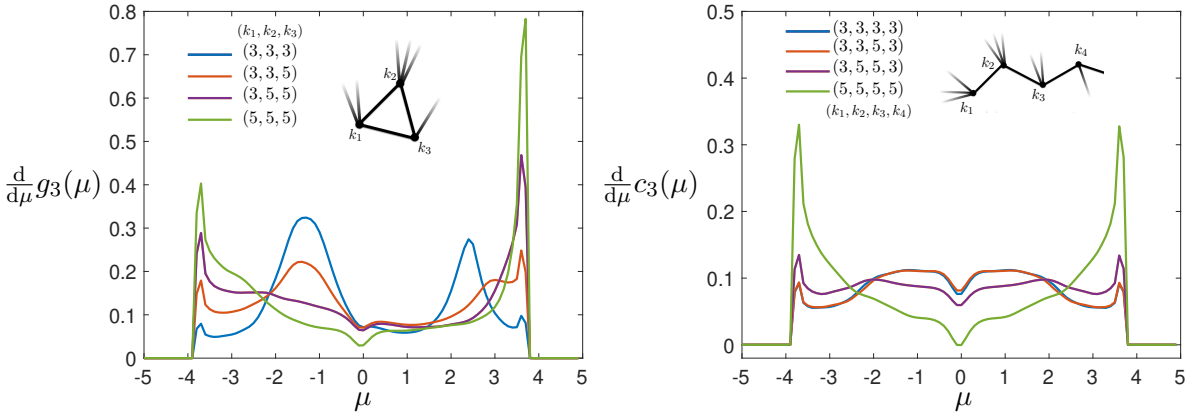


Figure 6.4: Spectral contributions from loops and chains on large graphs with degree distribution $p(k) = \frac{1}{2}(\delta_{k,3} + \delta_{k,5})$. On the left panel we show the average spectral contribution of a loop of length 3. The average, $\left\langle \frac{\mathrm{d}}{\mathrm{d}\mu} g_3(\mu) \right\rangle$, is shown separately for the four combinations of degrees an embedded triangle can have in this case. On the right panel we show the average, $\left\langle \frac{\mathrm{d}}{\mathrm{d}\mu} c_3(\mu) \right\rangle$, for four combinations of degrees. See (6.28) and (6.31) for definitions.

Table 6.1: In this table we show the mean and standard deviation of a sample of 1000 different of integral (6.48) and of (6.49). For each sample a different set $\{(k_p, \boldsymbol{x}_p)\}_{p=1,\dots,\ell}$ and $\{(k, \boldsymbol{x}), (k_p, \boldsymbol{x}_p)\}_{p=1,\dots,\ell}$ was drawn to calculate $g_\ell(\mu)$ and $c_\ell(\mu)$ respectively. The digit in parenthesis corresponds to the standard deviation and is of the order of the last digit to its left. The incredibly small variation shows that all samples individually are very close to the conjectured values of 2 for $g_3(\mu)$ and 0 for the rest.

Integral	$\ell = 3$	$\ell = 4$	$\ell = 5$	$\ell = 6$	$\ell = 7$
$\int \mathrm{d}\mu\, \mu^3 \frac{\mathrm{d}}{\mathrm{d}\mu} g_\ell(\mu)$	1.995(8)	0.0010(7)	-0.001(1)	0.0008(6)	0.0005(4)
$\int \mathrm{d}\mu\, \mu^3 \frac{\mathrm{d}}{\mathrm{d}\mu} c_\ell(\mu)$	0.0006(6)	0.0007(6)	0.0006(5)	0.0007(5)	0.0007(5)

number of closed paths around each node. Note that given that we interpret this loop as being infinitely far away from any other loop in the graph, we do not expect to be any other contributions into the moments from non backtracking paths from the rest of the graph except for the loop itself. This is because the rest of the graph has a tree-like topology. For odd ℓ's we observe asymmetric contributions as odd moments are related to the number of odd loops, this is the case as can be seen in Figure 6.4. For the case of the chain contributions, $c_\ell(\mu|\{x, x_p\}_{p=1,\dots,\ell})$ we expect all of its odd moments to vanish since the spectra of chains are symmetric, an example in Figure 6.4.

For the specific choice of $\hat{\varrho}(\mu) = \alpha\mu^3$ we can then conjecture that only the terms with $g_3(\mu)$ will contribute as no other integral will be non zero. That is, we expect that for *any*

sample of $W[\boldsymbol{x}]$ we will have,

$$\int \mathrm{d}\mu\ \mu^3 \frac{\mathrm{d}}{\mathrm{d}\mu} g_\ell(\mu|\{x_p\}_{p=1,\ldots,\ell}) = 2\delta_{3,\ell} \qquad \begin{array}{c}\text{when the } \boldsymbol{x}_p\text{'s are samples} \\ \text{of } f_{k_p-2}[\boldsymbol{x}_p]\end{array} . \tag{6.48}$$

Since no other term contains a loop of length 3.

By the same reasoning no open chain should have any triangle contribution either.

$$\int \mathrm{d}\mu\ \mu^3 \frac{\mathrm{d}}{\mathrm{d}\mu} c_\ell(\mu|x, \{x_p\}_{p=1,\ldots,\ell}) = 0 \qquad \begin{array}{c}\text{when } \boldsymbol{x} \text{ and the } \boldsymbol{x}_p\text{'s are sampled} \\ \text{from } f_{k-2}[\boldsymbol{x}] \text{ and } f_{k-1}[\boldsymbol{x}] \text{ according to (6.31)}\end{array} \tag{6.49}$$

This has also been tested extensively numerically and the results agree. In Table 6.1 we show the empirical averages over 10^3 samples from the population for the integrals for ℓ up to 7. Note that the statistics reported in Table 6.1 are for different draws of $\{(k_p, \boldsymbol{x}_p)\}_{p=1,\ldots,\ell}$ and $\{(k, \boldsymbol{x}), (k_p, \boldsymbol{x}_p)\}_{p=1,\ldots,\ell}$. This is important as conjectures (6.48) and (6.49) are *for each instance* in particular and not results that hold only on average. The small standard deviations observed in Table 6.1, of at most $8 \times 10^{-}3$, reflect the fact that almost all elements where in fact very close to the conjectured values.

We would like to point out that to calculate these integrals, (6.48) and (6.49), accurately the correct value for $g_\ell(\mu)$ and $c_\ell(\mu)$ had to be used, according to definitions (6.28) and (6.31). This means choosing the correct branch of *log* or otherwise getting the incorrect value of the integral. Numerically this means calculating $\mathrm{TrLog}(\mathbf{A})$ and not $\mathrm{Log}\det(\mathbf{A})$, confirming for this case the choice discussed in Chapter 4, (4.50)

We can then evaluate the generating function by using the fact that $W[\boldsymbol{x}|\alpha\mu^3] = W[\boldsymbol{x}|0]$ and that for the $\mathcal{O}(1/N)$ contributions (6.48) and (6.49) imply that

$$\begin{aligned} R(\{k_p\}_{p=1,\ldots,\ell}) =& 2\alpha\delta_{3,\ell} \\ C(k, \{k_p\}_{p=1,\ldots,\ell}) =& 0. \end{aligned} \tag{6.50}$$

Additionally, we also note that (6.48) and (6.49) imply that $\mathcal{Z}_k = 1$, then the effective degree distribution becomes

$$\tilde{p}(k) = p(k)\frac{k}{c}, \tag{6.51}$$

First we note that the generating function can be written in the following way (with a slight abuse of notation for the functional dependence), for the case $\hat{\varrho}(\mu) = 0$ all the $\hat{\varrho}$ dependent terms in the leading order of (6.27) cancel and we get,

$$\begin{aligned}\phi[0] =& \phi_0(N) + \frac{1}{N}\phi_1(N)\\ \phi_1(N) =& \lim\left[-\frac{1}{2}\mathrm{Tr}(\mathbf{B}) - \frac{1}{2}\mathrm{Tr}(\mathbf{MB}) + \frac{1}{4}\mathrm{Tr}(\mathbf{B}^2) + \sum_{\ell=3}^{\infty}\frac{1}{2\ell}\mathrm{Tr}((\mathbf{M}-\mathbf{B})^{\ell})\right]\Bigg|_{\forall \ell\, C_\ell, R_\ell = 0}\\ =& -\frac{1}{2}\left(\frac{\overline{k^2}}{c}\right) - \frac{1}{2}\left(\frac{\overline{k^2}}{c} - 1\right)\left(\frac{\overline{k^2}}{c}\right) + \frac{1}{4}\left(\frac{\overline{k^2}}{c}\right)^2 + \frac{1}{2}\left[-\log(2) + 1 - \frac{1}{2}\right]. \quad (6.52)\end{aligned}$$

Where we have used the fact that $\mathrm{Tr}(\mathbf{M}^p\mathbf{B}^q)|_{R_\ell, C_\ell = 0} = (\overline{k^2}/c - 1)^p(\overline{k^2}/c)^q$, the binomial expansion, and the logarithm series.

We now evaluate the generating function for $\hat{\varrho}(\mu) = \alpha\mu^3$, that is $\phi[\alpha\mu^3]$. In this approximation (6.27) of $\phi[\alpha\mu^3]$ differs from $\phi[0]$ by only on term, the one with $\mathrm{Tr}(\mathbf{M}^3)$. This is because according to (6.50), all the other terms have also have zero in the exponents. This leads to the following *familiar* results,

$$\begin{aligned}\phi(\alpha) =& \phi[\alpha\mu^3] = \phi[0] + \frac{1}{6N}\left(\frac{\overline{k^2}}{c} - 1\right)^3\left(\mathrm{e}^{6\alpha} - 1\right)\\ m(\alpha) =& \frac{\partial}{\partial\alpha}\phi(\alpha) = \frac{1}{N}\left(\frac{\overline{k^2}}{c} - 1\right)^3\mathrm{e}^{6\alpha}. \quad (6.53)\end{aligned}$$

We have thereby recovered exactly the results (3.16) and (3.17) from Chapter 3. These results were deduced by swapping the finite size triangle distribution of the CM with the known Poissonian asymptotic result, found in [5]. Amazingly we have deduced the same pair of formulae *without* using the rigorous result from [5], the replicas and the integrals have done all the work for us. It is incredible how integrals can recover such precise combinatorial results of discrete problems asymptotically, in this case the exact expected number of triangles in an infinitely large random graph.

At this point we are amazed that the replicas recovered a previous result, validating this formalism at least for the present problem. Much more can be done with a little extra effort. The replica method is not only good to recover the generating function by substituting a particular $\varrho(\mu)$, it can also help estimate the average spectral density by taking a functional derivative $\delta/\delta\hat{\varrho}(\mu)$, we will pursue this in the following section.

6.3.2 Formula for the spectrum for $\hat{\varrho}(\mu) = \alpha\mu^3$

To calculate a formula for the average spectral density we need to take the functional derivative w.r.t $\hat{\varrho}(\mu)$ of the generating function. As we have argued in the previous section, according to (6.47) we have in this case $\delta W/\delta\hat{\varrho}(\mu)|_{\hat{\varrho}(\mu)=\alpha\mu^3} = 0$, therefore we only need to look at explicit variations of the generating function w.r.t $\hat{\varrho}$. It is only a matter of looking at how the derivatives act on the two types of traces presented.

For the pure $\mathbf{M}$ traces,

$$\frac{\delta}{\delta\hat{\varrho}(\mu)}\lim \mathrm{Tr}(\mathbf{M}^{\ell}) = \left[\prod_{p=1}^{\ell}\sum_{k_p}\tilde{p}(k_p)(k_p-1)\right]\exp\left(\ell R(\{k_p\}_{p=1,\dots,\ell})\right) \times\left[\left\langle \ell\frac{\mathrm{d}}{\mathrm{d}\mu}g_{\ell}(\mu|\{x_p\}_{p=1,\dots,\ell})\right\rangle_{W_R} - \sum_{p=1}^{\ell}\left\langle\frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu)\right\rangle_{W_{k_p-1}}\right]. \quad (6.54)$$

Where we have introduced,

$$W_R[\{(k_p,\boldsymbol{x}_p)\}_{p=1,\dots,\ell}] = \frac{\prod_{p=1}^{\ell} f_{k_p-2}[\boldsymbol{x}_p]\ \cdot\ \mathrm{e}^{\ell\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}g_{\ell}(\mu)}}{\int\prod_{p=1}^{\ell}\mathcal{D}\boldsymbol{x}_p f_{k_p-2}[\boldsymbol{x}_p]\mathrm{e}^{\ell\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}g_{\ell}(\mu)}} = \prod_{p=1}^{\ell} f_{k_p-2}[\boldsymbol{x}_p]. \quad (6.55)$$

Note that the last equality is only valid because for $\hat{\varrho}(\mu) = \alpha\mu^3$ we are assuming (6.48), the first expression is more general. Examples of $\left\langle\frac{\mathrm{d}}{\mathrm{d}\mu}g_{\ell}(\mu|\{x_p\}_{p=1,\dots,\ell})\right\rangle_{W_R}$ can be seen on the left of Figure 6.4 for different degree combinations. Note that all loopy contributions come with a "penalty" for each node, $-\left\langle\frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu)\right\rangle_{W_k}$ that is also normalized. This is natural because corrections to the leading order that is already normalized should integrate to 0.

And for the other terms,

$$\frac{\delta}{\delta\hat{\varrho}(\mu)}\lim \mathrm{Tr}(\prod_{t=1}^{s}\mathbf{M}^{\ell_t-1}\mathbf{B}^{m_t+1}) = \left[\prod_{t=1}^{s}\prod_{p_t=1}^{\ell_t+m_t}\sum_{k_{p_t}}\tilde{p}(k_{p_t})\right]\prod_{t=1}^{s}\left[\prod_{p_t=1}^{\ell_t-1}(k_{p_t}-1)\prod_{p_t=\ell_t}^{\ell_t+m_t}k_{p_t}\right]$$
$$\times\exp\left(\sum_{t=1}^{s}\left[(\ell_t+1)C(k_{\ell_{t-1}+m_{t-1}},\{k_{p_t}\}_{p_t=1,\dots,\ell_t}) + \sum_{p_t=\ell_t+1}^{\ell_t+m_t}2C(k_{p_t-1},k_{p_t})\right]\right)$$
$$\times\left[\sum_{t=1}^{s}\left(\left\langle(\ell_t+1)\frac{\mathrm{d}}{\mathrm{d}\mu}c_{\ell}(\mu|x,\{x_p\}_{p=1,\dots,\ell})\right\rangle_{W_C} - \sum_{p_t=1}^{\ell_t}\left\langle\frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu)\right\rangle_{W_{k_{p_t}-1}} - \left\langle\frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu)\right\rangle_{W_{k_{\ell_{t-1}+m_{t-1}}}}\right.\right.$$
$$\left.\left. + \sum_{p_t=\ell_t+1}^{\ell_t+m_t}\left\langle 2\frac{\mathrm{d}}{\mathrm{d}\mu}c_1(\mu|x,x_1)\right\rangle_{W_C} - \left\langle\frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu)\right\rangle_{W_{k_{(p_t-1)}}} - \left\langle\frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu)\right\rangle_{W_{k_{p_t}-1}}\right)\right]. \quad (6.56)$$

With,

$$W_C[(k,\boldsymbol{x}),\{(k_p,\boldsymbol{x}_p)\}_{p=1,\dots,\ell}] = \frac{f_{k-1}[\boldsymbol{x}]f_{k_\ell-1}[\boldsymbol{x}_\ell]\prod_{p=1}^{\ell-1} f_{k_p-2}[\boldsymbol{x}_p]\,\mathrm{e}^{\ell\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}c_\ell(\mu|x,\{x_p\}_{p=1,\dots,\ell})}}{\mathcal{Z}_C}$$
$$= f_{k-1}[\boldsymbol{x}]f_{k_\ell-1}[\boldsymbol{x}_\ell]\prod_{p=1}^{\ell-1} f_{k_p-2}[\boldsymbol{x}_p]. \tag{6.57}$$

Where again,

$$\mathcal{Z}_C = \int \mathcal{D}\boldsymbol{x} f_{k-1}[\boldsymbol{x}]\mathcal{D}\boldsymbol{x}_\ell f_{k_\ell-1}[\boldsymbol{x}_\ell]\prod_{p=1}^{\ell-1}\mathcal{D}\boldsymbol{x}_p f_{k_p-2}[\boldsymbol{x}_p]\mathrm{e}^{\ell\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}c_\ell(\mu|x,\{x_p\}_{p=1,\dots,\ell})}$$
$$= 1. \tag{6.58}$$

Here the last equalities on the second lines of (6.57) and (6.58) are only valid for this particular case of $\hat{\varrho}(\mu) = \alpha\mu^3$ using (6.48) and (6.49). The first lines of (6.57) and (6.58) are general results. Examples of $\left\langle \frac{\mathrm{d}}{\mathrm{d}\mu}c_\ell(\mu|x,\{x_p\}_{p=1,\dots,\ell})\right\rangle_{W_C}$ can be seen in the right panel of Figure 6.4

We can now follow a similar reasoning to the previous section, we first look at what the result is for $\hat{\varrho}(\mu) = 0$

$$\varrho(\mu)|_{\hat{\varrho}=0} = \varrho_0(\mu) + \frac{1}{N}\varrho_1(\mu) \tag{6.59}$$
$$\varrho_0(\mu) = \lim_{N\to\infty}\frac{\delta\phi}{\delta\hat{\varrho}(\mu)}[0] \tag{6.60}$$
$$\varrho_1(\mu) = \left[\frac{\delta}{\delta\hat{\varrho}(\mu)}\lim\left\{-\frac{1}{2}\mathrm{Tr}(\mathbf{B}) - \frac{1}{2}\mathrm{Tr}(\mathbf{MB}) + \frac{1}{4}\mathrm{Tr}(\mathbf{B}^2) + \sum_{\ell=3}^{\infty}\frac{1}{2\ell}\mathrm{Tr}((\mathbf{M}-\mathbf{B})^\ell)\right\}\right]\Bigg|_{\hat{\varrho}=0}. \tag{6.61}$$

We have already proven in section 6.2.2 that $\varrho_0(\mu)$ is equal to (6.43), it is the known result for the spectral density of the CM. Contrary to (6.52), the subleading order $\varrho_1(\mu)$ in this case we expect to be a diverging series, as is the case for regular graphs $p(k) = \delta_{k,q}$, as was shown in the previous chapter. While we cannot offer an analytic resummation of the series analogous to (5.76), we can simply postulate its existence as we expect the $\mathcal{O}(1/N)$ corrections to the unbiased case to be finite, $\varrho_1(\mu) = \mathcal{O}(1)$.

Again, using (6.48) and (6.49) we can claim that for $\hat{\varrho}(\mu) = \alpha\mu^3$ the formula for the spectral density differs from the $\hat{\varrho}(\mu) = 0$ case only by the term with $\mathrm{Tr}(\mathbf{M}^3)$. This leads

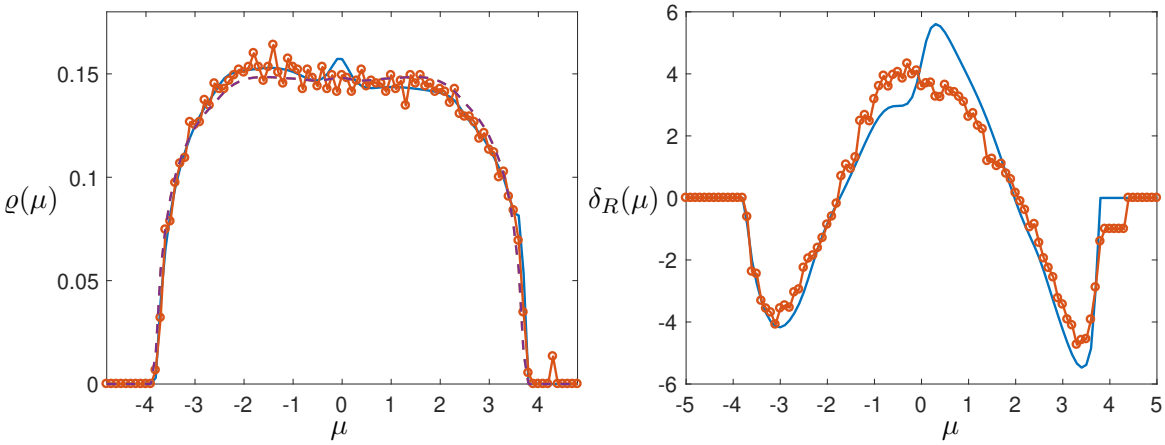


Figure 6.5: We compare theory versus numerical samples for the spectral density of loopy graphs with degree distribution $p(k) = \frac{1}{2}(\delta_{k,3} + \delta_{k,5})$. On the left panel we show the average spectral density of a sample of 10 graphs of $N = 750$ nodes with $\alpha = 0.35$ and $m = 0.1$. Orange circles correspond to the samples obtained via MCMC, the purple dashed line corresponds to the spectral density for the unbiased CM with the same degree distribution, $\varrho_0(\mu)$. The solid blue line corresponds to $\varrho_0(\mu)$ plus the loopy contribution $\tilde{\varrho}_1(\mu)$. On the right panel we plot $\delta_R(\mu)$, the scaled deviations of the CDF from $R_0(\mu)$ defined in (6.64). Orance circles correspond to the numerical samples and the solid blue line to the theoretical value, the integral of (6.63).

to

$$\varrho(\mu)|_{\hat{\varrho}(\mu)=\alpha\mu^3} = \varrho_0(\mu) + \frac{1}{N}\varrho_1(\mu) + \frac{1}{N}\tilde{\varrho}_1(\mu) \tag{6.62}$$

$$\begin{aligned}\tilde{\varrho}_1(\mu) =& \frac{1}{2}(\mathrm{e}^{6\alpha} - 1)\left[\prod_{p=1}^{3}\sum_{k_p}\tilde{p}(k_p)(k_p - 1)\right] \\ &\times\left[\left\langle \frac{\mathrm{d}}{\mathrm{d}\mu} g_\ell(\mu|\{x_p\}_{p=1,\dots,\ell})\right\rangle_{W_R} - \frac{1}{3}\sum_{p=1}^{3}\left\langle \frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu)\right\rangle_{W_{k_p-1}}\right].\end{aligned} \tag{6.63}$$

Note this is the generalization of (5.86) to an arbitrary degree distribution.

To test this expression we sampled loopy graphs from the ensemble with the MCMC method, choosing $p(k) = \frac{1}{2}(\delta_{k,3} + \delta_{k,5})$, $N = 750$, and $\alpha = 0.35$. For this values the expected loop density is $m = 0.37$ while the observed one is $m_{MCMC} = 0.36(4)$, one order of magnitude than the one expected for the CM in this case, which is $m_0 \approx 0.04$. This has a noticeable effect on the average spectral density, which should then be asymmetrical. This can be observed in Figure 6.5. We observe a good agreement between simulations and theory (6.62). In this case we are only comparing the average of 10 samples, nevertheless the agreement is decent enough. To better show the power of the theory we show the difference between CDF's which are less noisy. Since all the derivatives $\frac{\mathrm{d}}{\mathrm{d}\mu}$ in (6.62)

appear explicitly, we can integrate it analytically and therefore give an estimate for the following quantity.

$$\delta_R(\mu) = \int_{-\infty}^{\mu} \mathrm{d}\mu N(\varrho(\mu) - \varrho_0(\mu)) \approx \int_{-\infty}^{\mu} \mathrm{d}\mu\, \tilde{\varrho}_1(\mu). \tag{6.64}$$

Note we have discarded the typical fluctuations $\varrho_1(\mu)$ as they are one order of magnitude below for this value of α as pointed out before. In Figure 6.5 a good agreement can be seen. The differences are probably due to finite size effects of a higher order than $1/N$ or to the missing terms $\varrho_1(\mu)$ not calculated.

6.4 Discussion

We have presented a general approach for spectrally constrained ensembles with a prescribed degree sequence, generalizing the results of Chapter 5 for regular graphs. We demonstrated that it recovers the previous theory for the spectral density of degree constrained random graphs sampled uniformly, [39], while opening the door to new extensions. We show its validity for the nontrivial case of a triangle bias. We highlight that we were able to recover for (6.53) both the exponential factor $\mathrm{e}^{6\alpha}$ and the prefactor $\frac{1}{6}(\overline{k^2}/c - 1)^3$, which corresponds to the expected number of triangles in the CM. It is very interesting that we were able to recover a result from graph theory just by solving the integral of the imaginary replica formalism through the saddle point approximation. Additionally we point out we did not restrict ourselves to the conventional ER or regular cases, the result was derived for an arbitrary degree distribution $p(k)$. It is also interesting to observe that with the same formalism we were able to recover both the number of loops and the spectral contribution of each loop. A purely cavity like approach would probably be able to calculate the contribution of a single loop, like in [22], but it would not enable one to estimate the number of triangles, as it only deals with a particular instance and not with an ensemble. The number of loops in graph instances has to be added by hand.

So far, finite size corrections due to loops in uniform models on sparse graphs had only been tested for regular and ER graphs, for spin glasses, [64, 68], and for spectral calculations on regular graphs, [63]. Actually this series of papers inspired the author's calculations for both this chapter and the previous one, [48]. For the case of spin systems an scaling of the summands of the series in the $\mathcal{O}(1/N)$ terms was found, and thus it was possible to evaluate the finite size corrections without calculating an extremely large

amount of terms at an extreme computational cost. For the case of regular graph the series is divergent but is possible to give a finite answer thanks to exact analytic expression. For the case of an arbitrary degree distribution we would need to study how such complicated multiple integrals over distributions behave. So far we are not aware of any work were such a complicated diverging sum has been tamed. Nevertheless it was not necessary, as with the triangle bias we were able to single out a contribution, $\mathrm{Tr}(\mathbf{M}^3)$ in this case, and prove the formulas presented. We expect the same to be true for other kinds of bias, as long as the effect of the bias is of larger magnitude than the natural fluctuations of the ensemble. It would be interesting to attempt to do a resummation of a series like (6.33) or to find an argument for a scaling behavior for large ℓ.

The are many future directions to be explored. We aim to investigate if the theory recovers correctly the finite size corrections for other choices of $\hat{\varrho}(\mu)$ and even more if there exist $\hat{\varrho}(\mu)$ such that equations (6.26) describe a leading order which is different than the one for CM. So far it seems like the integral is only taking into account tree-like structures at leading order, nevertheless it could still include degree correlations, for example.

Another natural extension is to include weights in the graphs. This would add minor modifications if the weights are independent and identically distributed when $\hat{\varrho}(\mu) = 0$. Such extensions have been studied for a particular case of $\hat{\varrho}(\mu) = y[\theta(\mu - a) - \theta(\mu - b)]$ on weighted graphs on [55, 58]. There it was shown that indeed a bias on the leading order could be induced and the weighted populations dynamics described in section 6.2.1 was used successfully. This gives strong indication that at least the weighted version of this ensemble could work at leading order for a certain family of biases $\hat{\varrho}(\mu)$.

Chapter 7

Conclusions and outlook

In this thesis we have explored in depth a certain family of exponential random loopy graphs. Both theoretically and numerically. We believe that anyone who wishes to work with ensemble (3.4) now has a very good picture of the overall behavior of this ensemble. This should provide a good starting point for further research.

As the magnitude of the Lagrange parameter α is increased, the loopy ensemble (3.4) has a nice controllable first regime of weakly interacting loops and then quickly transitions into a clustered regime and eventually shatters into many small graphs. For anyone attempting to use it as a random graph null ensemble we have provided good analytic estimates for the first regime and a good estimate of where and why the transition occurs. The size dependence of the phenomenology of this ensemble is quite striking. Especially its entropic nature. It is quite counter-intuitive to note that the local moves, the edge swaps, will produce dramatically different graphs on initial conditions where the only difference is the number of nodes. This is an important lesson, as entropic transitions will probably be present in other random graph ensembles.

A possible fix to the shattering transition observed for ensemble (3.4) would be to add as an additional hard constraint the property of having only one giant component. This would allow to explore the space of loop connected graphs for larger graph sizes. While this would prevent the shattering transition to occur by definition, the clustering transition might still occur early on. To prevent the clustering transition from occurring so dramatically, one could also include a soft constraint on the interaction between loops $r(\mathbf{A})$. In order to play with these ideas numerically a careful revision of the MCMC algorithm should be done. As we have mentioned, sampling from (3.4) requires nontrivial acceptance probabilities, (A.5), in order to ensure correct sampling. Any modification of

the hard constraints would require a careful revision of the MCMC algorithm as well.

As mentioned before, in light of the phenomenology of (3.4), even though less straightforward to analyze, the spectrally constrained ensemble (5.3) seems like a good alternative. The functional imaginary replica approach was originally introduced to try to describe networks with an extensive number of loops, [47]. While we have not achieved that quite yet, it turned out to correctly describe ensembles with a triangle bias in the weak clustering regime. This is highly nontrivial as it was shown to be useful for modest graph sizes all the way down to $N = 750$. This provides an analytical tool with which to investigate the effect of a small amount of loopiness in processes determined by the spectral density, for example in the spherical model [53, 74]. Even though it has only been tested so far for a triangle bias, the enormous zoo of diagrams that emerged in the series expansion promises interesting contributions with direct topological interpretability when testing different kinds of biases. For example we could test which kind of subgraphs are favored when enforcing nontrivial spectral biases, at least at the level of $\mathcal{O}(1/N)$ corrections. This could range from biasing certain types of loops, to certain degree correlations.

Additionally, the rich structure of the solutions of the functional saddle point equation, shown in Figure 6.3, points towards possible interesting connections with other branches of abstract mathematics like algebraic geometry. Also, the fact that simpler theories on weighted random graphs [55] do show effects in the leading order makes the prospect of studying more complicated models of weighted sparse random graphs with the functional imaginary replica method promising. It is only a matter of investigating what is the interpretation of biasing a degree constrained model with respect to the moments of the weighted adjacency matrix. The most interesting result would be to find correlations between weights. We think that at least the functional approach provides a very robust and unified way of phrasing many different ensembles in a single set of equations, including $\mathcal{O}(1/N)$ corrections.

Finally, the spectrally constrained ensemble (4.5) can also be tested numerically, without worrying about the analytic solution. This is a very demanding task computationally. It requires an MCMC algorithm that diagonalizes the adjacency matrix for each proposed edge swap, as described in section 5.3.4 of Chapter 5. As a numerical experiment, we biased a 4 regular graph to increase the number of eigenvalues outside of the normal domain, $[-\sqrt{12}, \sqrt{12}]$, by choosing $\hat{\varrho}(\mu) = -200[\theta(\mu + \sqrt{12}) - \theta(\mu - \sqrt{12})]$. Preliminary results do in fact show the formation of loopy structures on a 4-regular graph 100 nodes without

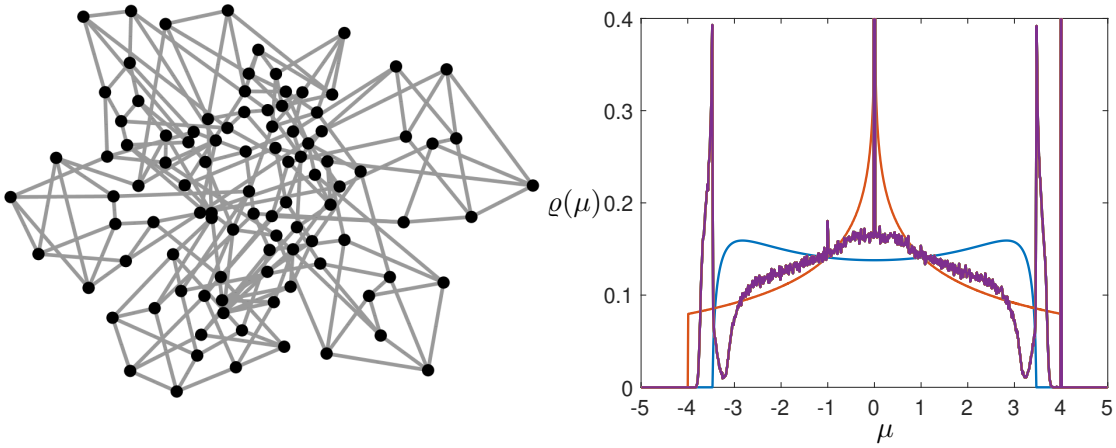


Figure 7.1: Results of sampling numerical via MCMC from (6.1) with a functional Lagrange parameter $\hat{\varrho}(\mu) = -200[\theta(\mu + \sqrt{12}) - \theta(\mu - \sqrt{12})]$ with $p(k) = \delta_{k,4}$ and $N = 100$. On the left a single graph obtained vi the MCMC showing a higher presence of squares, *i.e.* not locally tree-like. On the right comparison between the average spectral density of this ensemble (purple), 10^4 samples, the spectral density of a 4-regular random graph (blue), and the spectral density of a 2 dimensional lattice (orange).

breaking it into smaller graphs. This can be seen in Figure 7.1 where we show one example of a graph sampled from this ensemble. In this case the density of squares around each node is 6.4, close to the value 8 for a 2 dimensional lattice but still very far from 36 which would correspond to the shattered phase of square bias. We also show the average spectral density obtained through the MCMC sampling compared with that of random 4-regular graph and to that of a 2 dimensional lattice. A closer resemblance to the spectral density of a 2 dimensional lattice could be achieved by adding further biases in a larger amount of intervals. These preliminary results suggest that it could be a good strategy to simply randomize a real network with edge swaps while trying to keep the same spectral density as that of the observed network in order to generate realistic looking graphs. This shows that it is a promising direction to develop further analytic theory of spectrally constrained ensembles beyond the one presented in this thesis constrained to finite size corrections of $\mathcal{O}(1/N)$.

Our ensemble can be generalized in many natural ways, including either loops of longer length, degree correlations, other spectral observables, etc. Nevertheless the question remains what properties should be included to generate more realistic looking random networks. It is hard to imagine that there is a single observable or set of observables that upon tuning will make the ensemble mimic real networks very accurately. In fact it

is unlikely that such a universal ensemble exists, for the reason that real networks come from many different contexts. Networks themselves are a projection of a more complicated object, an abstraction or simplification of reality. It is conceivable that to generate realistic looking graphs one needs to consider their particular field of origin.

It is safe to say the subject of random loopy graphs is far from being solved. Not only due to the technical difficulty of tackling this problem analytically but also due to the nontrivial behavior of such models. We hope this thesis provides important steps in further developments of this very interesting subject.

Bibliography

1. Euler, L. Solutio problematis ad geometriam situs pertinentis. *Commentarii academiae scientiarum Petropolitanae,* 128–140 (1741).
2. Casella, G. & Berger, R. L. *Stat. Inference* (Duxbury Pacific Grove, CA, 2002).
3. Solomonoff, R. & Rapoport, A. Connectivity of random nets. *B. Math. Biophys.* **13,** 107–117 (1951).
4. Erdös, P. & Rényi, A. On random graphs, I. *Publicationes Mathematicae (Debrecen)* **6,** 290–297 (1959).
5. Bollobás, B. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Eur. J. Combin.* **1,** 311–316 (1980).
6. Annibale, A., Roberts, E. & Coolen, A. *Generating Random Networks and Graphs* (Oxford University Press, 2017).
7. Newman, M. E. The structure and function of complex networks. *SIAM review* **45,** 167–256 (2003).
8. Cover, T. M. & Thomas, J. A. *Elements of information theory* (John Wiley & Sons, 2012).
9. Coolen, A., De Martino, A. & Annibale, A. Constrained Markovian dynamics of random graphs. *J. Stat. Phys.* **136,** 1035–1067 (2009).
10. Strauss, D. On a general class of models for interaction. *SIAM review* **28,** 513–527 (1986).
11. Burda, Z., Jurkiewicz, J. & Krzywicki, A. Network transitivity and matrix models. *Phys. Rev. E* **69,** 026106 (2004).
12. Jonasson, J. The random triangle model. *J. Appl. Probab.* **36,** 852–867 (1999).

13. Chatterjee, S., Diaconis, P., *et al.* Estimating and understanding exponential random graph models. *Ann. Stat.* **41,** 2428–2461 (2013).

14. Yin, M. A detailed investigation into near degenerate exponential random graphs. *J. Stat. Phys.* **164,** 241–253 (2016).

15. Newman, M. E. Random graphs with clustering. *Phys. Rev. Lett.* **103,** 058701 (2009).

16. Miller, J. C. Percolation and epidemics in random clustered networks. *Phys. Rev. E* **80,** 020901 (2009).

17. Hackett, A., Melnik, S. & Gleeson, J. P. Cascades on a class of clustered random networks. *Phys. Rev. E* **83,** 056107 (2011).

18. Volz, E. M., Miller, J. C., Galvani, A. & Meyers, L. A. Effects of heterogeneous and clustered contact patterns on infectious disease dynamics. *PLoS Comput. Biol.* **7,** e1002042 (2011).

19. Herrero, C. P. Ising model in clustered scale-free networks. *Phys. Rev. E* **91,** 052812 (2015).

20. Peron, T. K. D., Ji, P., Kurths, J. & Rodrigues, F. A. Spectra of random networks in the weak clustering regime. *Europhys. Lett.* **121,** 68001 (2018).

21. Herrero, C. P. Self-avoiding walks and connective constants in clustered scale-free networks. *Phys. Rev. E* **99,** 012314 (2019).

22. Cantwell, G. T. & Newman, M. Message passing on networks with loops. *P. Natl. Acad. Sci. USA* **116,** 23398–23403 (2019).

23. Heath, L. S. & Parikh, N. Generating random graphs with tunable clustering coefficients. *Physica A* **390,** 4577–4587 (2011).

24. Foster, D., Foster, J., Paczuski, M. & Grassberger, P. Communities, clustering phase transitions, and hysteresis: Pitfalls in constructing network ensembles. *Phys. Rev. E* **81,** 046115 (2010).

25. Tamm, M., Shkarin, A., Avetisov, V., Valba, O. & Nechaev, S. Islands of stability in motif distributions of random networks. *Phys. Rev. Lett.* **113,** 095701 (2014).

26. Avetisov, V. *et al.* Eigenvalue tunneling and decay of quenched random network. *Phys. Rev. E* **94,** 062313 (2016).

27. López, F. A., Barucca, P., Fekom, M. & Coolen, A. C. Exactly solvable random graph ensemble with extensively many short cycles. *J. Phys. A: Math. Theor.* **51,** 085101 (2018).

28. Avetisov, V., Gorsky, A., Maslov, S., Nechaev, S. & Valba, O. Phase transitions in social networks inspired by the Schelling model. *Phys. Rev. E* **98,** 032308 (2018).

29. Pospelov, N. *et al.* Spectral peculiarity and criticality of a human connectome. *Phys. of life reviews* **31,** 240–256 (2019).

30. Avetisov, V., Gorsky, A., Nechaev, S. & Valba, O. Localization and non-ergodicity in clustered random networks. *J. of Complex Networks* **8,** cnz026 (2020).

31. Fedoryuk, M. *The saddle-point method* (Nauka, Moscow, 1977).

32. Livan, G., Novaes, M. & Vivo, P. *Introduction to random matrices: theory and practice* (Springer, 2018).

33. Rogers, T., Castillo, I. P., Kühn, R. & Takeda, K. Cavity approach to the spectral density of sparse symmetric random matrices. *Phys. Rev. E* **78,** 031116 (2008).

34. Metz, F. L., Neri, I. & Bollé, D. Spectra of sparse regular graphs with loops. *Phys. Rev. E* **84,** 055101 (2011).

35. Sokhotskii, Y. W. On definite integrals and functions used in series expansions. *St. Petersburg* (1873).

36. Edwards, S. F. & Jones, R. C. The eigenvalue spectrum of a large symmetric random matrix. *J. Phys. A: Math. Gen.* **9,** 1595 (1976).

37. Kühn, R., Van Mourik, J., Weigt, M. & Zippelius, A. Finitely coordinated models for low-temperature phases of amorphous systems. *J. Phys. A: Math. Theor.* **40,** 9227 (2007).

38. Mézard, M., Parisi, G. & Virasoro, M. *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications* (World Scientific Publishing Company, 1987).

39. Kühn, R. Spectra of sparse random matrices. *J. Phys. A: Math. Theor.* **41,** 295002 (2008).

40. Kühn, R. & Van Mourik, J. Spectra of modular and small-world matrices. *J. Phys. A: Math. Theor.* **44,** 165205 (2011).

41. Barucca, P. Spectral density of equitable core–periphery graphs. *Physica A,* 124649 (2020).

42. Verbaarschot, J. & Zirnbauer, M. Critique of the replica trick. *J. Phys. A: Math. Gen.* **18,** 1093 (1985).

43. Zirnbauer, M. R. Another critique of the replica trick. *arXiv preprint cond-mat/9903338* (1999).

44. Fyodorov, Y. V. & Mirlin, A. On the density of states of sparse random matrices. *J. Phys. A: Math. Gen.* **24,** 2219 (1991).

45. Lin, T.-F. Problem of the disordered chain. *J. Math. Phys.* **11,** 1584–1590 (1970).

46. McKay, B. D. The expected eigenvalue distribution of a large regular graph. *Linear Algebra Appl.* **40,** 203–216 (1981).

47. Coolen, A. *Replica methods for loopy sparse random graphs* in *J. Phys. Conf. Ser.* **699** (2016), 012022.

48. López, F. A. & Coolen, A. C. Imaginary replica analysis of loopy regular random graphs. *J. Phys. A: Math. Theor.* **53,** 065002 (2020).

49. Alon, N. & Spencer, J. H. *The probabilistic method* (John Wiley & Sons, 2004).

50. Wormald, N. C. *et al.* Models of random regular graphs. *London Mathematical Society Lecture Note Series,* 239–298 (1999).

51. Bollobás, B. *Random graphs* **73** (Cambridge university press, 2001).

52. Roberts, E. & Coolen, A. Random graph ensembles with many short loops. *ESAIM: Proc. and Surveys* **47,** 97–115 (2014).

53. Berlin, T. H. & Kac, M. The spherical model of a ferromagnet. *Phys. Rev.* **86,** 821 (1952).

54. Metz, F. L. & Stariolo, D. A. Index statistical properties of sparse random graphs. *Phys. Rev. E* **92,** 042153 (2015).

55. Metz, F. L. & Castillo, I. P. Large deviation function for the number of eigenvalues of sparse random graphs inside an interval. *Phys. Rev. Lett.* **117,** 104101 (2016).

56. Metz, F. L. & Castillo, I. P. Level compressibility for the Anderson model on regular random graphs and the eigenvalue statistics in the extended phase. *Phys. Rev. B* **96,** 064202 (2017).

57. Castillo, I. P. & Metz, F. L. Large-deviation theory for diluted Wishart random matrices. *Phys. Rev. E* **97,** 032124 (2018).

58. Castillo, I. P. & Metz, F. L. Theory for the conditioned spectral density of noninvariant random matrices. *Phys. Rev. E* **98,** 020102 (2018).

59. Rodgers, G. J. & Bray, A. J. Density of states of a sparse random matrix. *Phys. Rev. B* **37,** 3557 (1988).

60. Kamenev, A. & Mézard, M. Wigner-Dyson statistics from the replica method. *J. Phys. A: Math. Gen.* **32,** 4373 (1999).

61. Cavagna, A., Garrahan, J. P. & Giardina, I. Index distribution of random matrices with an application to disordered systems. *Phys. Rev. B* **61,** 3960 (2000).

62. Metz, F. L. Replica-symmetric approach to the typical eigenvalue fluctuations of Gaussian random matrices. *J. Phys. A: Math. Theor.* **50,** 495002 (2017).

63. Metz, F. L., Parisi, G. & Leuzzi, L. Finite-size corrections to the spectrum of regular random graphs: An analytical solution. *Phys. Rev. E* **90,** 052109 (2014).

64. Lucibello, C., Morone, F., Parisi, G., Ricci-Tersenghi, F. & Rizzo, T. Finite-size corrections to disordered Ising models on random regular graphs. *Phys. Rev. E* **90,** 012146 (2014).

65. Khorunzhy, O., Shcherbina, M. & Vengerovsky, V. Eigenvalue distribution of large weighted random graphs. *J. Math. Phys.* **45,** 1648–1672 (2004).

66. Johnson, T. Exchangeable pairs, switchings, and random regular graphs. *arXiv preprint arXiv:1112.0704* (2011).

67. Kesten, H. Symmetric random walks on groups. *T. Am. Math. Soc.* **92,** 336–354 (1959).

68. Ferrari, U. *et al.* Finite-size corrections to disordered systems on Erdös-Rényi random graphs. *Phys. Rev. B* **88,** 184201 (2013).

69. Wormald, N. C. The asymptotic distribution of short cycles in random regular graphs. *J. Comb. Theory B* **31,** 168–182 (1981).

70. Gradshteyn, I. S. & Ryzhik, I. M. *Table of integrals, series, and products* (Academic press, 2014).

71. Gao, Z. & Wormald, N. C. Distribution of subgraphs of random regular graphs. *Random Struct. Algor.* **32,** 38–48 (2008).

72. Lucibello, C., Morone, F., Parisi, G., Ricci-Tersenghi, F. & Rizzo, T. Anomalous finite size corrections in random field models. *J. Stat. Mech.: Theory E.* **2014,** P10025 (2014).

73. Wall, H. S. *Analytic theory of continued fractions* (Courier Dover Publications, 2018).

74. Kosterlitz, J. M., Thouless, D. J. & Jones, R. C. Spherical model of a spin-glass. *Phys. Rev. Lett.* **36,** 1217 (1976).

75. Eggleton, R. B. & Holton, D. A. in *Combinatorial Mathematics VIII* 155–172 (Springer, 1981).

Appendix A

Numerical sampling of graphs via MCMC edge swap dynamics

In order for this thesis to be self-contained, we will present a brief recap of the algorithms described in [6, 9] used to generate the samples.

The main task of such algorithms is to define a Markov chain with the following characteristics.

$$p_{t+1}(\mathbf{A}) = \sum_{\mathbf{A}' \in \Omega_{\mathcal{M}}} W(\mathbf{A}|\mathbf{A}') p_t(\mathbf{A}'), \tag{A.1}$$

1. The measure p_t converges to the invariant measure $p_\infty(\mathbf{A}) = \frac{1}{Z} \mathrm{e}^{-H(\mathbf{A})}$.

2. The allowed transitions come from a limited set Φ of elementary moves

$$F : \Omega_F \subseteq \Omega_{\mathcal{M}} \to \Omega_{\mathcal{M}}$$

3. For each $F \in \Phi$ there exists a unique inverse F^{-1} that acts on the same set of graphs, $\Omega_{F^{-1}} = \Omega_F$

With these conditions we will be able to define a dynamical process that will allow us to sample effectively from ensemble (5.6). The reason we need nontrivial moves is to be sure we respect the degree constraint, a single edge dynamics cannot achieve this. The move we choose for this case is called an edge swap. It is done by choosing a pair of non-touching edges and interchanging them, see Figure A.1.

We now need to define the transition probabilities. They are chosen to obey detailed

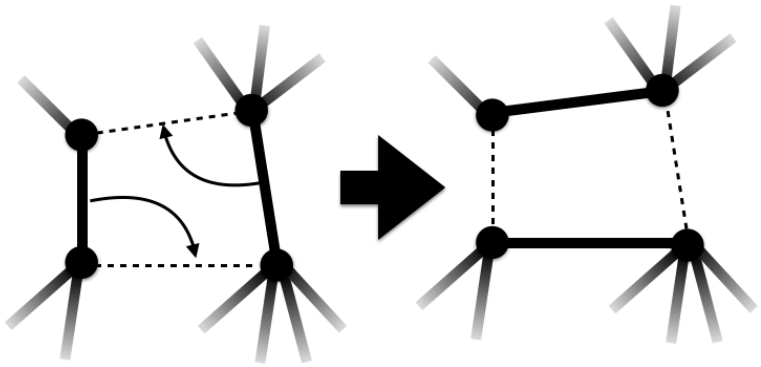


Figure A.1: Edge swap for MCMC dynamics in the space of simple nondirected graphs. This is the simplest type of move that leaves all degrees invariant.

balance with (5.6), $W(\mathbf{A}|\mathbf{A}')p_\infty(\mathbf{A}') = W(\mathbf{A}'|\mathbf{A})p_\infty(\mathbf{A})$, as a necessary condition to satisfy (i). Therefore we write the transition probabilities as

$$W(\mathbf{A}|\mathbf{A}') = \sum_{F\in\Omega'} \frac{I_F(\mathbf{A}')}{n(\mathbf{A}')} \left[\delta_{\mathbf{A},F\mathbf{A}'} A(F\mathbf{A}'|\mathbf{A}') + \delta_{\mathbf{A},\mathbf{A}'}\left[1 - A(F\mathbf{A}'|\mathbf{A}')\right]\right]. \tag{A.2}$$

Where we use the following definitions,

$$\begin{aligned}
\Omega' &= \{F \in \Phi | \exists \mathbf{A} \in \Omega_{\mathcal{M}} \text{ s. t. } F\mathbf{A} \neq \mathbf{A}\} \\
I_F(\mathbf{A}) &= \begin{cases} 1 & \text{if } \mathbf{A} \to F\mathbf{A} \text{ is an allowed move} \\ 0 & \text{otherwise} \end{cases} \\
n(\mathbf{A}) &= \sum_{F\in\Omega'} I_F(\mathbf{A}) \\
A(F\mathbf{A}|\mathbf{A}) &: \text{ acceptance probability of move } \mathbf{A} \to F\mathbf{A}
\end{aligned} \tag{A.3}$$

The interpretation of these transition probabilities is as follows. At each step a move is chosen uniformly at random from all possible moves. this is done with probability $1/n(\mathbf{A})$. After a move is chosen, it is accepted or rejected with probability $A(F\mathbf{A}|\mathbf{A})$. We fix the value of $A(F\mathbf{A}|\mathbf{A})$ by enforcing detailed balance

$$(\forall \mathbf{A} \in \Omega)(\forall F \in \Omega') : \; A(F\mathbf{A}|\mathbf{A})\mathrm{e}^{-H(\mathbf{A})}/n(\mathbf{A}) = A(\mathbf{A}|F\mathbf{A})\mathrm{e}^{-H(F\mathbf{A})}/n(F\mathbf{A}) \tag{A.4}$$

There are different choices that enforce this condition, but the one we take is the following

one.

$$A(\mathbf{A}|\mathbf{A}') = \frac{n(\mathbf{A}')\mathrm{e}^{-\frac{1}{2}[H(\mathbf{A})-H(\mathbf{A}')]}}{n(\mathbf{A}')\mathrm{e}^{-\frac{1}{2}[H(\mathbf{A})-H(\mathbf{A}')]} + n(\mathbf{A})\mathrm{e}^{-\frac{1}{2}[H(\mathbf{A}')-H(\mathbf{A})]}} = \frac{1}{1+\mathrm{e}^{E(\mathbf{A})-E(\mathbf{A}')}} \tag{A.5}$$

Here we have defined an effective energy $E(\mathbf{A}) = H(\mathbf{A}) + \log n(\mathbf{A})$ to stress the fact that the acceptance probability depends not only on the particular function $H(\mathbf{A})$ but also on the current state through $n(\mathbf{A})$ for any particular ensemble. In [6] it is shown that $n(\mathbf{A})$ an be written explicitly as

$$n(\mathbf{A}) = \frac{1}{4}\left(\sum_i k_i\right)^2 + \frac{1}{4}\sum_i k_i - \frac{1}{2}\sum_i k_i^2 - \frac{1}{2}\sum_{ij} k_i A_{ij} k_j + \frac{1}{4}\mathrm{Tr}\left(\mathbf{A}^4\right) + \frac{1}{2}\mathrm{Tr}\left(\mathbf{A}^3\right) \tag{A.6}$$

The ergodicity of this process is proven in [75].

Appendix B

Entropic argument for shattering transition

We define T and K to be the random variables corresponding to the number of isolated triangles in a graph, $T(\mathbf{A})$, and the number of cliques of degree q in a graph, $K_q(\mathbf{A})$, in the uniform configuration model respectively. Their distributions are:

$$\begin{aligned} P_N(T) =& \frac{\sum_{\mathbf{A}} \delta_{T,T(\mathbf{A})} \prod_{i=1}^{N} \delta_{k_i,\sum_j A_{ij}}}{\sum_{\mathbf{A}} \prod_{i=1}^{N} \delta_{k_i,\sum_j A_{ij}}}, \\ Q_N(K) =& \frac{\sum_{\mathbf{A}} \delta_{K,K_q(\mathbf{A})} \prod_{i=1}^{N} \delta_{k_i,\sum_j A_{ij}}}{\sum_{\mathbf{A}} \prod_{i=1}^{N} \delta_{k_i,\sum_j A_{ij}}}. \end{aligned} \tag{B.1}$$

Let us assume that both T and K are Poissonian random variables.

$$\begin{aligned} P_N(T) =& Poiss(T|\lambda_t) \\ Q_N(K) =& Poiss(K|c_q/N^{d-1}) \end{aligned} \tag{B.2}$$

We choose $d-1$ for consistency with the definition in the main body of the text. Since P_N and Q_N are probabilities of the uniform distribution over the degree constrained $\mathbf{A}$'s (CM), they are simply related to the number of graphs with that given number of triangles or cliques,

$$\frac{P_N(T)}{Q_N(K)} = \frac{\sum_{\mathbf{A}} \delta_{T,T(\mathbf{A})} \prod_{i=1}^{N} \delta_{k_i,\sum_j A_{ij}}}{\sum_{\mathbf{A}} \delta_{K,K(\mathbf{A})} \prod_{i=1}^{N} \delta_{k_i,\sum_j A_{ij}}} = \mathrm{e}^{-\lambda_t + \frac{c_q}{N^d}} \frac{(\lambda_t)^T}{(c_q/N^{d-1})^K} \frac{K!}{T!}. \tag{B.3}$$

Our question is, for a given loop density, are there more graphs that realize it asymptotically through triangles or through cliques? For this we need to write down the number of triangles and cliques in terms of the desired loop density. This gives the next two different values,

$$\begin{aligned} T =& \frac{m}{6}N \\ K =& \frac{m}{q(q^2-1)}N \end{aligned} \tag{B.4}$$

Then we look at the asymptotic limit

$$\begin{aligned} \lim_{N\to\infty} \frac{P_N\left(\frac{m}{6}N\right)}{Q_N\left(\frac{m}{q(q^2-1)}N\right)} = \lim_{N\to\infty} \exp\Bigg(&-\lambda_t + \frac{c_q}{N^d} + \frac{m}{6}N\log\lambda_t - \frac{m}{q(q^2-1)}N\log(c_q) \\ &+\frac{md}{q(q^2-1)}N\log(N) + \left(\frac{m}{q(q^2-1)}N\right)! - \left(\frac{m}{6}N\right)!\Bigg) \end{aligned} \tag{B.5}$$

Note that it is dominated by the $N\log N$ term, since $d = \frac{1}{2}q(q-1)$.

$$\lim_{N\to\infty} \frac{P_N\left(\frac{m}{6}N\right)}{Q_N\left(\frac{m}{q(q^2-1)}N\right)} = \lim_{N\to\infty} \exp\left(-m\left(\frac{1}{6} - \frac{1}{2(q+1)}\right)N\log N\right) = 0. \tag{B.6}$$

This means that for a fixed m and sufficiently large N there are always more graphs that achieve such loop density by creating isolated cliques than by triangles embedded in a larger component.

Appendix C

Appendices for the case of regular loopy graphs

In this appendix we include additional material necessary for Chapter 5.

C.1 The functional integral

C.1.1 Transformation to Fourier components

In this Appendix we simplify expression (5.20,5.21) for our generating function $\phi[\hat{\varrho}]$. We first introduce a number of definitions to compactify our formulae. We begin by introducing the Fourier coefficients of our functional order parameters:

$$P(\boldsymbol{\varphi},\omega)=\sum_{\ell\in\mathbb{Z}}W_\ell(\boldsymbol{\varphi})\frac{\mathrm{e}^{\mathrm{i}\ell\omega}}{2\pi},\qquad W_\ell(\boldsymbol{\varphi})=\int_{-\pi}^{\pi}\mathrm{d}\omega\ P(\boldsymbol{\varphi},\omega)\mathrm{e}^{-\mathrm{i}\ell\omega},\tag{C.1}$$

$$\hat{P}(\boldsymbol{\varphi},\omega)=\sum_{\ell\in\mathbb{Z}}\hat{W}_\ell(\boldsymbol{\varphi})\mathrm{e}^{-\mathrm{i}\ell\omega},\qquad \hat{W}_\ell(\boldsymbol{\varphi})=\int_{-\pi}^{\pi}\frac{\mathrm{d}\omega}{2\pi}\ \hat{P}(\boldsymbol{\varphi},\omega)\mathrm{e}^{\mathrm{i}\ell\omega}.\tag{C.2}$$

We then write $S[P,\hat{P}]$ strictly in terms of the Fourier transforms $\{W_\ell(\boldsymbol{\varphi}),\hat{W}_\ell(\boldsymbol{\varphi})\}$, noting that $\omega\in[-\pi,\pi]$. The result is

$$S[P,\hat{P}]=S[W_1,\hat{W}_1,W_2,\hat{W}_2]+S_0[\{W_\ell,\hat{W}_\ell\}],\tag{C.3}$$

where, using the notation $\mathbf{r} = \{r_\ell \geq 0,\ \ell \in \mathbb{Z}\}$,

$$\begin{aligned} S_0[\{W_\ell, \hat{W}_\ell\}] &= \mathrm{i} \sum_{\ell \notin \{1,2\}} \int \mathrm{d}\boldsymbol{\varphi}\ W_\ell(\boldsymbol{\varphi}) \hat{W}_\ell(\boldsymbol{\varphi}) \\ &\quad + \log \int \mathrm{d}\boldsymbol{\varphi}\ \nu(\boldsymbol{\varphi}) \sum_{\mathbf{r}} \delta_{q, \sum_{\ell \in \mathbb{Z}} \ell r_\ell} \prod_{\ell \in \mathbb{Z}} \frac{[-\mathrm{i}\hat{W}_\ell(\boldsymbol{\varphi})]^{r_\ell}}{r_\ell!}, \end{aligned} \tag{C.4}$$

$$\begin{aligned} S[W_1, \hat{W}_1, W_2, \hat{W}_2] &= -\frac{q}{2}(1+\frac{q-2}{2N}) + \mathrm{i} \int \mathrm{d}\boldsymbol{\varphi} \Big[W_1(\boldsymbol{\varphi}) \hat{W}_1(\boldsymbol{\varphi}) + W_2(\boldsymbol{\varphi}) \hat{W}_2(\boldsymbol{\varphi}) \Big] \\ &\quad - \frac{q}{2N} \int \mathrm{d}\boldsymbol{\varphi}\ W_2(\boldsymbol{\varphi}) \mathrm{e}^{-\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}} + \frac{q}{2}(1+\frac{q}{N}) \int \mathrm{d}\boldsymbol{\varphi} \mathrm{d}\boldsymbol{\varphi}'\ W_1(\boldsymbol{\varphi}) W_1(\boldsymbol{\varphi}') \mathrm{e}^{-\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'} \\ &\quad - \frac{q^2}{4N} \int \mathrm{d}\boldsymbol{\varphi} \mathrm{d}\boldsymbol{\varphi}'\ W_2(\boldsymbol{\varphi}) W_2(\boldsymbol{\varphi}') \mathrm{e}^{-2\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'}. \end{aligned} \tag{C.5}$$

The integration in (5.20) can be replaced by integration over the functional Fourier components, since (apart from an irrelevant multiplicative constant) the Jacobian of this coordinate transformation is unitary. We define $\mathcal{D}W = \prod_{\boldsymbol{\varphi}} [\mathrm{d}W(\boldsymbol{\varphi}) \sqrt{N\Delta_\varphi / 2\pi}]$, and the functional delta distribution $\delta[W] = \prod_{\boldsymbol{\varphi}} [\delta(W(\boldsymbol{\varphi})) \sqrt{2\pi / N\Delta_\varphi}]$, where $\delta(x)$ is the ordinary delta distribution, so that for any smooth $F[W]$ we will have:

$$\int \mathcal{D}W\ F[W]\delta[W] = F[0], \qquad \delta[W] = \int \mathcal{D}\hat{W}\ \mathrm{e}^{\mathrm{i}N \int \mathrm{d}\boldsymbol{\varphi}\ W(\boldsymbol{\varphi}) \hat{W}(\boldsymbol{\varphi})}, \tag{C.6}$$

and the generating function can be written as, modulo an irrelevant additive constant:

$$\phi[\hat{\varrho}] = \lim \frac{1}{N} \log \int \Big[\prod_{\ell \in \mathbb{Z}} \mathcal{D}W_\ell \mathcal{D}\hat{W}_\ell \Big] \mathrm{e}^{NS[W_1, \hat{W}_1, W_2, \hat{W}_2] + NS_0[\{W_\ell, \hat{W}_\ell\}]}. \tag{C.7}$$

With (A.6), integration over the Fourier components with $\ell = 1, 2$ has become trivial:

$$\begin{aligned} &\int \prod_{\ell \notin \{1,2\}} [\mathcal{D}W_\ell \mathcal{D}\hat{W}_\ell]\ \mathrm{e}^{NS_0[\{W_\ell, \hat{W}_\ell\}]} = \mathrm{e}^{\mathrm{i}N \sum_{\ell=1}^2 \int \mathrm{d}\boldsymbol{\varphi}\ W_\ell(\boldsymbol{\varphi}) \hat{W}_\ell(\boldsymbol{\varphi})} \\ &\quad \times \int \prod_{\ell \notin \{1,2\}} \big[\mathcal{D}\hat{W}_\ell\ \delta[\hat{W}_\ell]\big] \exp\Big(N \log \int \mathrm{d}\boldsymbol{\varphi}\ \nu(\boldsymbol{\varphi}) \sum_{\mathbf{r}} \delta_{q, \sum_{\ell \in \mathbb{Z}} \ell r_\ell} \prod_{\ell \in \mathbb{Z}} \frac{[-\mathrm{i}\hat{W}_\ell(\boldsymbol{\varphi})]^{r_\ell}}{r_\ell!} \Big) \\ &\quad = \exp\Big(\mathrm{i}N \sum_{\ell=1}^2 \int \mathrm{d}\boldsymbol{\varphi}\ W_\ell(\boldsymbol{\varphi}) \hat{W}_\ell(\boldsymbol{\varphi}) \\ &\qquad + N \log \int \mathrm{d}\boldsymbol{\varphi}\ \nu(\boldsymbol{\varphi}) \sum_{0 \leq r \leq q/2} \frac{[-\mathrm{i}\hat{W}_1(\boldsymbol{\varphi})]^{q-2r}}{(q-2r)!} \frac{[-\mathrm{i}\hat{W}_2(\boldsymbol{\varphi})]^r}{r!} \Big) \Big]. \end{aligned} \tag{C.8}$$

Hence (5.20) can be written as follows:

$$\begin{aligned}\phi[\hat{\varrho}] &= -\frac{q}{2}(1+\frac{q-2}{2N}) + \lim \frac{1}{N}\log\int \mathcal{D}\hat{W}_1\mathcal{D}\hat{W}_2 \exp\Big\{N\Big(S_1[\hat{W}_1]+S_2[\hat{W}_2] \\ &\quad + N\log\int \mathrm{d}\boldsymbol{\varphi}\ \nu(\boldsymbol{\varphi}) \sum_{0\leq r\leq q/2} \frac{[-\mathrm{i}\hat{W}_1(\boldsymbol{\varphi})]^{q-2r}}{(q-2r)!}\frac{[-\mathrm{i}\hat{W}_2(\boldsymbol{\varphi})]^{r}}{r!}\Big)\Big\}, \end{aligned} \tag{C.9}$$

with the following two functionals $S_{1,2}[\hat{W}]$:

$$\begin{aligned} S_1[\hat{W}] &= \frac{1}{N}\log\int \mathcal{D}W \exp\Big(\mathrm{i}N\int \mathrm{d}\boldsymbol{\varphi}\ W(\boldsymbol{\varphi})\hat{W}(\boldsymbol{\varphi}) \\ &\qquad +\frac{q}{2}(N+q)\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\ W(\boldsymbol{\varphi})U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}')\Big) \end{aligned} \tag{C.10}$$

$$\begin{aligned} S_2[\hat{W}] &= \frac{1}{N}\log\int \mathcal{D}W \exp\Big(\mathrm{i}N\int \mathrm{d}\boldsymbol{\varphi}\ W(\boldsymbol{\varphi})\hat{W}(\boldsymbol{\varphi}) - \frac{q}{2}\int \mathrm{d}\boldsymbol{\varphi}\ W(\boldsymbol{\varphi})V(\boldsymbol{\varphi}) \\ &\qquad -\frac{1}{4}q^2\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\ W(\boldsymbol{\varphi})U_2(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}')\Big), \end{aligned} \tag{C.11}$$

where we used the short-hands $U_n(\boldsymbol{\varphi},\boldsymbol{\varphi}') = \exp(-\mathrm{i}n\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}')$ and $V(\boldsymbol{\varphi}) = U_1(\boldsymbol{\varphi},\boldsymbol{\varphi})$.

C.1.2 Gaussian functional integrals

Both $S_1[\hat{W}]$ and $S_2[\hat{W}]$ involve complex Gaussian functional integrals, of the following form, with $d = \dim(\boldsymbol{\varphi})$ and $\Delta_\varphi \to 0$ in the functional integration limit:

$$\begin{aligned} J[U,Q] &= \int \mathcal{D}W\ \mathrm{e}^{-\frac{1}{2}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\ W(\boldsymbol{\varphi})U(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}')+\int \mathrm{d}\boldsymbol{\varphi}\ W(\boldsymbol{\varphi})Q(\boldsymbol{\varphi})} \\ &= \int\prod_{\boldsymbol{\varphi}}\frac{\mathrm{d}W(\boldsymbol{\varphi})}{\sqrt{2\pi/N\Delta_\varphi}}\ \mathrm{e}^{-\frac{1}{2}\Delta_\varphi^2\sum_{\boldsymbol{\varphi}\boldsymbol{\varphi}'}W(\boldsymbol{\varphi})U(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}')+\Delta_\varphi\sum_{\boldsymbol{\varphi}}W(\boldsymbol{\varphi})Q(\boldsymbol{\varphi})} \\ &= \Big(\frac{N}{\Delta_\varphi}\Big)^{d/2}\frac{1}{\sqrt{\mathrm{Det}\mathbf{U}}}\ \mathrm{e}^{\frac{1}{2}\Delta_\varphi^2\sum_{\boldsymbol{\varphi}\boldsymbol{\varphi}'}(\Delta_\varphi^2\mathbf{U})^{-1}(\boldsymbol{\varphi},\boldsymbol{\varphi}')Q(\boldsymbol{\varphi})Q(\boldsymbol{\varphi}')}. \end{aligned} \tag{C.12}$$

$\mathbf{U}$ is the matrix of discretized values $U(\boldsymbol{\varphi},\boldsymbol{\varphi}')$. The entries of the inverse functional kernel $U^{-1}(\boldsymbol{\varphi},\boldsymbol{\varphi}')$, defined by the condition $\delta(\boldsymbol{\varphi}-\boldsymbol{\varphi}') = \int \mathrm{d}\boldsymbol{\varphi}''\ U^{-1}(\boldsymbol{\varphi},\boldsymbol{\varphi}'')U(\boldsymbol{\varphi}'',\boldsymbol{\varphi}')$, are $U^{-1}(\boldsymbol{\varphi},\boldsymbol{\varphi}') = (\Delta_\varphi^2\mathbf{U})^{-1}(\boldsymbol{\varphi},\boldsymbol{\varphi}')$. Hence the following identities hold:

$$J[U,Q] = \Big(\frac{N}{\Delta_\varphi}\Big)^{d/2}\frac{1}{\sqrt{\mathrm{Det}\mathbf{U}}}\ \mathrm{e}^{\frac{1}{2}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'Q(\boldsymbol{\varphi})U^{-1}(\boldsymbol{\varphi},\boldsymbol{\varphi}')Q(\boldsymbol{\varphi}')}, \tag{C.13}$$

$$J[U^{-1},Q] = (N\Delta_\varphi)^{d/2}\ \sqrt{\mathrm{Det}\mathbf{U}}\ \mathrm{e}^{\frac{1}{2}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'Q(\boldsymbol{\varphi})U(\boldsymbol{\varphi},\boldsymbol{\varphi}')Q(\boldsymbol{\varphi}')}. \tag{C.14}$$

We can now work out (C.10) and (C.11), and find

$$\begin{aligned}
\mathrm{e}^{NS_1[\hat{W}]} &= J\big[-q(N+q)U_1, \mathrm{i}N\hat{W}\big] \\
&= \Big(\frac{N}{-q(N+q)\Delta_\varphi}\Big)^{d/2}\frac{1}{\sqrt{\mathrm{Det}\mathbf{U}_1}}\ \mathrm{e}^{\frac{N^2}{2q(N+q)}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\ \hat{W}(\boldsymbol{\varphi})U_1^{-1}(\boldsymbol{\varphi},\boldsymbol{\varphi}')\hat{W}(\boldsymbol{\varphi}')}, \qquad \text{(C.15)} \\
\mathrm{e}^{NS_2[\hat{W}]} &= J\big[\frac{1}{2}q^2U_2, \mathrm{i}N\hat{W}-\frac{1}{2}qV\big] \\
&= \Big(\frac{2N}{q^2\Delta_\varphi}\Big)^{d/2}\frac{1}{\sqrt{\mathrm{Det}\mathbf{U}_2}}\ \mathrm{e}^{-\frac{1}{q^2}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'[N\hat{W}(\boldsymbol{\varphi})+\frac{q}{2}\mathrm{i}V(\boldsymbol{\varphi})]U_2^{-1}(\boldsymbol{\varphi},\boldsymbol{\varphi}')[N\hat{W}(\boldsymbol{\varphi}')+\frac{q}{2}\mathrm{i}V(\boldsymbol{\varphi}')]} \qquad \text{(C.16)}
\end{aligned}$$

For (C.9) this implies:

$$\begin{aligned}
\phi[\hat{\varrho}] &= -\frac{q}{2}(1+\frac{q-2}{2N}) + \lim\frac{1}{N}\log\Big(\frac{2N}{q^2\Delta_\varphi}\Big)^{d/2}\frac{1}{\sqrt{\mathrm{Det}\mathbf{U}_2}}\int\mathcal{D}\hat{W}_1\ \mathrm{e}^{NS_1[\hat{W}_1]} \\
&\quad\times\int\mathcal{D}\hat{W}_2\ \mathrm{e}^{-\frac{1}{q^2}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'[N\hat{W}_2(\boldsymbol{\varphi})+\frac{q}{2}\mathrm{i}V(\boldsymbol{\varphi})]U_2^{-1}(\boldsymbol{\varphi},\boldsymbol{\varphi}')[N\hat{W}_2(\boldsymbol{\varphi}')+\frac{q}{2}\mathrm{i}V(\boldsymbol{\varphi}')]} \\
&\quad\times\exp\Big\{N\log\int\mathrm{d}\boldsymbol{\varphi}\ \nu(\boldsymbol{\varphi})\sum_{0\le r\le q/2}\frac{[-\mathrm{i}\hat{W}_1(\boldsymbol{\varphi})]^{q-2r}}{(q-2r)!}\frac{[-\mathrm{i}\hat{W}_2(\boldsymbol{\varphi})]^r}{r!}\Big\}. \qquad \text{(C.17)}
\end{aligned}$$

We transform $\hat{W}_2(\boldsymbol{\varphi}) \to \hat{W}_2(\boldsymbol{\varphi})/N - q\mathrm{i}V(\boldsymbol{\varphi})/2N$ and expand for large N, following [63]. [1] This assures that $\phi[\hat{\varrho}]$ remains well-defined and nontrivial, and that the leading orders in N of the generating function can be written as follows (assuming that $q > 1$):

$$\begin{aligned}
\phi[\hat{\varrho}] =& -\frac{q}{2}(1+\frac{q-2}{2N}) + \lim\frac{1}{N}\log\Big(\frac{2}{Nq^2\Delta_\varphi}\Big)^{d/2}\frac{1}{\sqrt{\mathrm{Det}\mathbf{U}_2}}\int\mathcal{D}\hat{W}_1\ \mathrm{e}^{NS_1[\hat{W}_1]} \\
&\times\mathrm{e}^{N\log\int\mathrm{d}\boldsymbol{\varphi}\ \nu(\boldsymbol{\varphi})[-\mathrm{i}\hat{W}_1(\boldsymbol{\varphi})]^q/q!}\int\mathcal{D}\hat{W}_2\ \mathrm{e}^{-\frac{1}{2}\int\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\ \hat{W}_2(\boldsymbol{\varphi})(\frac{1}{2}q^2U_2)^{-1}(\boldsymbol{\varphi},\boldsymbol{\varphi}')\hat{W}_2(\boldsymbol{\varphi}')} \\
&\times\exp\Big\{-\mathrm{i}q(q-1)\frac{\int\mathrm{d}\boldsymbol{\varphi}\ \nu(\boldsymbol{\varphi})[-\mathrm{i}\hat{W}_1(\boldsymbol{\varphi})]^{q-2}[\hat{W}_2(\boldsymbol{\varphi})-\frac{1}{2}\mathrm{i}qV(\boldsymbol{\varphi})]}{\int\mathrm{d}\boldsymbol{\varphi}'\ \nu(\boldsymbol{\varphi}')[-\mathrm{i}\hat{W}_1(\boldsymbol{\varphi}')]^q} + \mathcal{O}(\frac{1}{N})\Big\}. \qquad \text{(C.18)}
\end{aligned}$$

We can now integrate over $\hat{W}_2$, and with the short-hand

$$R[\hat{W}(\boldsymbol{\varphi})] = q(q-1)\frac{\nu(\boldsymbol{\varphi})[-\mathrm{i}\hat{W}(\boldsymbol{\varphi})]^{q-2}}{\int\mathrm{d}\boldsymbol{\varphi}'\ \nu(\boldsymbol{\varphi}')[-\mathrm{i}\hat{W}(\boldsymbol{\varphi}')]^q} \qquad \text{(C.19)}$$

[1] Note that we could also have chosen the transformation $\hat{W}_2(\boldsymbol{\varphi}) \to \hat{W}_2(\boldsymbol{\varphi})/\sqrt{N} - q\mathrm{i}\mathrm{e}^{-\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}}/2N$, but this would in subsequent stages of the calculation have prompted a further rescaling of $\hat{W}_2$ by $\sqrt{N}$.

the result takes the form

$$\begin{aligned}
\phi[\hat{\varrho}] &= -\frac{q}{2}(1+\frac{q-2}{2N}) + \lim \frac{1}{N} \log \int \mathcal{D}\hat{W}_1 \ \mathrm{e}^{-\frac{1}{4}q^2 \int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}' R[\hat{W}_1(\boldsymbol{\varphi})]U_2(\boldsymbol{\varphi},\boldsymbol{\varphi}')R[\hat{W}_1(\boldsymbol{\varphi}')]} \\
&\quad \times \mathrm{e}^{NS_1[\hat{W}_1]+N\log\int \mathrm{d}\boldsymbol{\varphi}\ \nu(\boldsymbol{\varphi})[-\mathrm{i}\hat{W}_1(\boldsymbol{\varphi})]^q/q!-\frac{1}{2}\int \mathrm{d}\boldsymbol{\varphi}\ R[\hat{W}_1(\boldsymbol{\varphi})]V(\boldsymbol{\varphi})+\mathcal{O}(N^{-1})} \\
&= -\frac{q}{2}(1+\frac{q-2}{2N}) + \lim \frac{1}{N} \log \int \mathcal{D}\hat{W}_1 \ \mathrm{e}^{-\frac{1}{4}q^2 \int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}' R[\hat{W}_1(\boldsymbol{\varphi})]U_2(\boldsymbol{\varphi},\boldsymbol{\varphi}')R[\hat{W}_1(\boldsymbol{\varphi}')]} \\
&\quad \times \Big(\frac{N}{-q(N+q)\Delta_{\varphi}}\Big)^{d/2} \frac{1}{\sqrt{\mathrm{Det}\mathbf{U}_1}} \ \mathrm{e}^{\frac{N^2}{2q(N+q)}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\ \hat{W}_1(\boldsymbol{\varphi})U_1^{-1}(\boldsymbol{\varphi},\boldsymbol{\varphi}')\hat{W}_1(\boldsymbol{\varphi}')} \\
&\quad \times \mathrm{e}^{N\log\int \mathrm{d}\boldsymbol{\varphi}\ \nu(\boldsymbol{\varphi})[-\mathrm{i}\hat{W}_1(\boldsymbol{\varphi})]^q/q!-\frac{1}{2}q\int \mathrm{d}\boldsymbol{\varphi}\ R[\hat{W}_1(\boldsymbol{\varphi})]V(\boldsymbol{\varphi})+\mathcal{O}(N^{-1})}. \qquad \text{(C.20)}
\end{aligned}$$

Finally we transform $\hat{W}_1(\boldsymbol{\varphi}) = \mathrm{i}\int \mathrm{d}\boldsymbol{\varphi}'\ U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}')$, which gives apart from irrelevant additive constants:

$$\phi[\hat{\varrho}] = \lim \frac{1}{N} \log \Big\{ \sqrt{\mathrm{Det}(q\mathbf{U}_1)} \int \mathcal{D}W \ \mathrm{e}^{N\mathcal{S}_0[W]+\mathcal{S}_1[W]+\mathcal{O}(N^{-1})} \Big\}, \qquad \text{(C.21)}$$

with

$$\begin{aligned}
\mathcal{S}_0[W] &= -\frac{q}{2}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\ W(\boldsymbol{\varphi})U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}') \\
&\quad + \log \int \mathrm{d}\boldsymbol{\varphi}\ \nu(\boldsymbol{\varphi}) \Big[\int \mathrm{d}\boldsymbol{\varphi}'\ U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}') \Big]^q, \qquad \text{(C.22)} \\
\mathcal{S}_1[W] &= -\frac{1}{4}(q-1)^2 \int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\ r[W(\boldsymbol{\varphi})]U_2(\boldsymbol{\varphi},\boldsymbol{\varphi}')r[W(\boldsymbol{\varphi}')] \qquad \text{(C.23)} \\
&\quad + \frac{1}{2}q^2 \int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\ W(\boldsymbol{\varphi})U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}') - \frac{1}{2}(q-1)\int \mathrm{d}\boldsymbol{\varphi}\ r[W(\boldsymbol{\varphi})]V(\boldsymbol{\varphi}),
\end{aligned}$$

and

$$r[W(\boldsymbol{\varphi})] = \frac{\nu(\boldsymbol{\varphi})[\int \mathrm{d}\boldsymbol{\varphi}'\ U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}')]^{q-2}}{\int \mathrm{d}\boldsymbol{\varphi}'\ \nu(\boldsymbol{\varphi}')[\int \mathrm{d}\boldsymbol{\varphi}''\ U_1(\boldsymbol{\varphi}',\boldsymbol{\varphi}'')W(\boldsymbol{\varphi}'')]^q}. \qquad \text{(C.24)}$$

C.1.3 Leading two orders via saddle point integration

Expression (C.21) allows us in the usual manner to calculate the leading two orders in N of the generating function, by substituting $W = W_0 + N^{-\frac{1}{2}}W_1$, where W_0 is the saddle point of $\mathcal{S}_0[W]$ and where $W_1 = \mathcal{O}(1)$. We obtain, again modulo a constant:

$$\phi[\hat{\varrho}] = \lim \Big\{ \mathcal{S}_0[W_0] + \frac{1}{N}\mathcal{S}_1[W_0] + \frac{1}{N}\log\Big[\frac{\sqrt{\mathrm{Det}(q\mathbf{U}_1)}}{\sqrt{\mathrm{Det}(-\boldsymbol{\Gamma})}}\Big]\Big\} + \mathcal{O}(N^{-\frac{3}{2}}), \qquad \text{(C.25)}$$

in which we have the functional curvature at the saddle point:

$$\Gamma(\boldsymbol{\varphi},\boldsymbol{\varphi}') = \frac{\delta^2\mathcal{S}[W]}{\delta W(\boldsymbol{\varphi})\delta W(\boldsymbol{\varphi}')}\Big|_{W_0}. \tag{C.26}$$

What remains is to compute $W_0(\boldsymbol{\varphi})$ and $\Gamma(\boldsymbol{\varphi},\boldsymbol{\varphi}')$. Setting $\delta\mathcal{S}_0/\delta W = 0$, and using the symmetry of U_1, gives the saddle point equation from which to solve W_0:

$$W_0(\boldsymbol{\varphi}) = \frac{1}{Z_q}\nu(\boldsymbol{\varphi})\Big[\int \mathrm{d}\boldsymbol{\varphi}'\; U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W_0(\boldsymbol{\varphi}')\Big]^{q-1}, \tag{C.27}$$

$$Z_q = \int \mathrm{d}\boldsymbol{\varphi}\; \nu(\boldsymbol{\varphi})\Big[\int \mathrm{d}\boldsymbol{\varphi}'\; U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W_0(\boldsymbol{\varphi}')\Big]^{q}, \tag{C.28}$$

and the curvature at the saddle point is found to be

$$\begin{aligned}
\Gamma(\boldsymbol{\varphi},\boldsymbol{\varphi}') &= \frac{q(q-1)}{Z_q}\int \mathrm{d}\boldsymbol{\psi}\; \nu(\boldsymbol{\psi})U_1(\boldsymbol{\varphi},\boldsymbol{\psi})U_1(\boldsymbol{\psi},\boldsymbol{\varphi}')\Big[\int \mathrm{d}\boldsymbol{\psi}'\; U_1(\boldsymbol{\psi},\boldsymbol{\psi}')W_0(\boldsymbol{\psi}')\Big]^{q-2} \\
&\quad -qU_1(\boldsymbol{\varphi},\boldsymbol{\varphi}') - \frac{q^2}{Z_q^2}\int \mathrm{d}\boldsymbol{\psi}\; \nu(\boldsymbol{\psi})U_1(\boldsymbol{\psi},\boldsymbol{\varphi})\Big[\int \mathrm{d}\boldsymbol{\psi}'\; U_1(\boldsymbol{\psi},\boldsymbol{\psi}')W_0(\boldsymbol{\psi}')\Big]^{q-1} \\
&\qquad \times \int \mathrm{d}\boldsymbol{\psi}\; \nu(\boldsymbol{\psi})U_1(\boldsymbol{\psi},\boldsymbol{\varphi}')\Big[\int \mathrm{d}\boldsymbol{\psi}'\; U_1(\boldsymbol{\psi},\boldsymbol{\psi}')W_0(\boldsymbol{\psi}')\Big]^{q-1} \\
&= q(q-1)\int \mathrm{d}\boldsymbol{\psi}\; U_1(\boldsymbol{\varphi},\boldsymbol{\psi})r[W_0(\boldsymbol{\psi})]U_1(\boldsymbol{\psi},\boldsymbol{\varphi}') - qU_1(\boldsymbol{\varphi},\boldsymbol{\varphi}') \\
&\qquad -q^2\Big[\int \mathrm{d}\boldsymbol{\psi}\; U_1(\boldsymbol{\varphi},\boldsymbol{\psi})W_0(\boldsymbol{\psi})\Big]\Big[\int \mathrm{d}\boldsymbol{\psi}\; U_1(\boldsymbol{\varphi}',\boldsymbol{\psi})W_0(\boldsymbol{\psi})\Big] \\
&= -q\int \mathrm{d}\boldsymbol{\psi}\; U_1(\boldsymbol{\phi},\boldsymbol{\psi})\Big[\delta(\boldsymbol{\psi}-\boldsymbol{\varphi}') - T(\boldsymbol{\psi},\boldsymbol{\varphi}')\Big],
\end{aligned} \tag{C.29}$$

with

$$T(\boldsymbol{\varphi},\boldsymbol{\varphi}') = (q-1)r[W_0(\boldsymbol{\varphi})]U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}') - qW_0(\boldsymbol{\varphi})\int \mathrm{d}\boldsymbol{\varphi}'' U_1(\boldsymbol{\varphi}',\boldsymbol{\varphi}'')W_0(\boldsymbol{\varphi}''). \tag{C.30}$$

We could also have written the curvature in the form $\mathbf{\Gamma} = -q\mathbf{U}^{\frac{1}{2}}(\mathbf{I}-\mathbf{T})\mathbf{U}_1^{\frac{1}{2}}$ with a symmetric kernel $\mathbf{T}$, but since we only require the determinant of $\mathbf{\Gamma}$ this would not make a difference. Various terms in $\phi[\hat{\rho}]$ can be simplified using equations (C.27,C.28). For instance, with the simple identity $\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}' W_0(\boldsymbol{\varphi})U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W_0(\boldsymbol{\varphi}') = 1$ we find that

$$\mathcal{S}_0[W_0] = -\frac{q}{2} + \log Z_q, \tag{C.31}$$

$$\begin{aligned}
\mathcal{S}_1[W_0] &= \frac{1}{2}q^2 - \frac{1}{4}(q-1)^2\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\; r[W(\boldsymbol{\varphi})]U_2(\boldsymbol{\varphi},\boldsymbol{\varphi}')r[W(\boldsymbol{\varphi}')] \\
&\qquad -\frac{1}{2}(q-1)\int \mathrm{d}\boldsymbol{\varphi}\; r[W(\boldsymbol{\varphi})]V(\boldsymbol{\varphi}).
\end{aligned} \tag{C.32}$$

Hence, using $\log \mathrm{Det}\mathbf{A} = \mathrm{Tr}\log\mathbf{A}$ and apart from irrelevant additive constants:

$$\begin{aligned}\phi[\hat{\varrho}] &= \lim\Big\{\log Z_q - \frac{(q-1)^2}{4N}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\; r[W_0(\boldsymbol{\varphi})]U_2(\boldsymbol{\varphi},\boldsymbol{\varphi}')r[W_0(\boldsymbol{\varphi}')] \\ &\quad -\frac{q-1}{2N}\int \mathrm{d}\boldsymbol{\varphi}\; r[W_0(\boldsymbol{\varphi})]V(\boldsymbol{\varphi}) - \frac{1}{2N}\mathrm{Tr}\log(\mathbf{I}-\mathbf{T})\Big\} + \mathcal{O}(N^{-\frac{3}{2}}). \end{aligned} \tag{C.33}$$

We can expand the last nontrivial term using $\mathrm{Tr}\log(\mathbf{I}-\mathbf{T}) = -\sum_{\ell=1}^{\infty}\mathrm{Tr}(\mathbf{T}^\ell)/\ell$. The first two terms in this sum give, after some simple manipulations:

$$\mathrm{Tr}(\mathbf{T}) = \int \mathrm{d}\boldsymbol{\varphi}\; T(\boldsymbol{\varphi},\boldsymbol{\varphi}) = (q-1)\int \mathrm{d}\boldsymbol{\varphi}\; r[W_0(\boldsymbol{\varphi})]V(\boldsymbol{\varphi}) - q, \tag{C.34}$$

$$\begin{aligned}\mathrm{Tr}(\mathbf{T}^2) &= \int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\; T(\boldsymbol{\varphi},\boldsymbol{\varphi}')T(\boldsymbol{\varphi}',\boldsymbol{\varphi}) \\ &= 2q - q^2 + (q-1)^2\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\; r[W_0(\boldsymbol{\varphi})]U_2(\boldsymbol{\varphi},\boldsymbol{\varphi}')r[W_0(\boldsymbol{\varphi}')]. \end{aligned} \tag{C.35}$$

Thus the first two terms $\ell = 1, 2$ precisely remove those non-constant terms in $\phi[\hat{\varrho}]$ that originated from $\mathcal{S}_1[W_0]$. This simplifies $\phi[\hat{\varrho}]$ to the following expression, modulo additive constants and $\mathcal{O}(N^{-\frac{3}{2}})$ terms, and upon inserting the definition of Z_q:

$$\phi[\hat{\varrho}] = \lim\Big\{\log\int \mathrm{d}\boldsymbol{\varphi}\;\nu(\boldsymbol{\varphi})\Big[\int \mathrm{d}\boldsymbol{\varphi}'\; U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W_0(\boldsymbol{\varphi}')\Big]^q + \frac{1}{2N}\sum_{\ell=3}^{\infty}\frac{\mathrm{Tr}(\mathbf{T}^\ell)}{\ell}\Big\}. \tag{C.36}$$

C.2 Replica symmetric value of the traces

Here we compute the traces $\mathrm{Tr}(\mathbf{M}^\ell)$, that appear in the generating function $\phi[\hat{\varrho}]$, for the kernel (5.42). Upon defining $\boldsymbol{\mu}^\star = \boldsymbol{\mu} + (q-2)(\mathbf{X}^*)^{-1}$ we can write this kernel as

$$M(\boldsymbol{\varphi},\boldsymbol{\varphi}') = \mathcal{Z}_{q-1}^{-1}\mathrm{e}^{\frac{1}{2}\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\mu}^\star\boldsymbol{\varphi} - \mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'}. \tag{C.37}$$

We can write the ℓ-th trace of $\mathbf{M}$ as follows, with the identification $\boldsymbol{\varphi}_{\ell+1} \equiv \boldsymbol{\varphi}_1$, and that both $\boldsymbol{\mu}$ and $\mathbf{X}^\star$ (and its inverse) are diagonal matrices in the space of $\boldsymbol{\varphi}$:

$$\begin{aligned}
\mathrm{Tr}(\mathbf{M}^\ell) &= \mathcal{Z}_{q-1}^{-\ell} \int \Big(\prod_{k=1}^{\ell} \mathrm{d}\boldsymbol{\varphi}_k \Big) \Big(\prod_{k=1}^{\ell} \mathrm{e}^{\frac{1}{2}\mathrm{i}\boldsymbol{\varphi}_k \cdot \boldsymbol{\mu}^\star \boldsymbol{\varphi}_k - \mathrm{i}\boldsymbol{\varphi}_k \cdot \boldsymbol{\varphi}_{k+1}} \Big) \\
&= \mathcal{Z}_{q-1}^{-\ell} \int \Big(\prod_{k=1}^{\ell} \mathrm{d}\boldsymbol{\varphi}_k \Big) \mathrm{e}^{\frac{1}{2}\mathrm{i} \sum_{kk'=1}^{\ell} \boldsymbol{\varphi}_k \cdot \left[\boldsymbol{\mu}^\star \delta_{kk'} - (\delta_{k+1,k'} + \delta_{k-1,k'}) \mathbf{I} \right] \boldsymbol{\varphi}_{k'}} \\
&= \mathcal{Z}_{q-1}^{-\ell} \prod_{\mu} \Bigg\{ \prod_{\alpha_\mu=1}^{n_\mu} \Big[\int \Big(\prod_{k=1}^{\ell} \mathrm{d}\phi_k \Big) \mathrm{e}^{-\frac{\mathrm{i}}{2} \sum_{kk'=1}^{\ell} \phi_k \left(\delta_{k,k'+1} + \delta_{k,k'-1} - \delta_{kk'} \mu^\star \right) \phi_{k'}} \Big] \\
&\qquad \times \prod_{\beta_\mu=1}^{m_\mu} \Big[\int \Big(\prod_{k=1}^{\ell} \mathrm{d}\psi_k \Big) \mathrm{e}^{\frac{\mathrm{i}}{2} \sum_{kk=1'}^{\ell} \psi_k \left(\delta_{k,k'+1} + \delta_{k,k'-1} - \delta_{kk'} \overline{\mu^\star} \right) \psi_{k'}} \Big] \Bigg\} \\
&= \mathcal{Z}_{q-1}^{-\ell} \prod_{\mu} \Big[Z(\mu_\varepsilon | \mathbf{A}^\star_{\ell,\mu})^{n_\mu} \overline{Z(\mu_\varepsilon | \mathbf{A}^\star_{\ell,\mu})}^{m_\mu} \Big]. \qquad \text{(C.38)}
\end{aligned}$$

Here $Z(\mu_\varepsilon | \mathbf{A}^\star_{\ell,\mu})$ denotes the original complex Gaussian integral defined in (5.10), and $\mathbf{A}^\star_{\ell,\mu}$ is now the adjacency matrix of a loop of length ℓ in the presence of a complex field acting on the diagonal, of value $(2-q)/x^\star(\mu)$:

$$\big(\mathbf{A}^\star_{\ell,\mu} \big)_{kk'} = \delta_{k,k'+1} + \delta_{k,k'-1} + \frac{2-q}{x^\star(\mu)} \delta_{kk'} \qquad (\text{with } k \bmod \ell). \qquad \text{(C.39)}$$

C.3 Recovering the Kesten-MacKay law

Here we derive expressions (5.51) and (5.52). The factor between curly brackets in the first line of the generating function (5.50), which will give the eigenvalue spectrum for graphs from the ensemble (5.6) in the limit $N \to \infty$, is given by the following expression, in which $x^\star(\mu)$ is given by (5.38):

$$\begin{aligned}
\varrho_0(\mu) &= \frac{1}{2\pi} \frac{\mathrm{d}}{\mathrm{d}\mu} \, \mathrm{Im} \Big[(q-2) \log \Big(x^\star(\mu) - \frac{1}{x^\star(\mu)} \Big) - q \log x^\star(\mu) \Big] \\
&= \frac{1}{2\pi} \frac{\mathrm{d}}{\mathrm{d}\mu} \Big[(q-2) \mathrm{Arg}\big(x^\star(\mu) - 1 \big) + (q-2) \mathrm{Arg}\big(x^\star(\mu) + 1 \big) - 2(q-1) \mathrm{Arg}\big(x^\star(\mu) \big) \Big]. \qquad \text{(C.40)}
\end{aligned}$$

We note that as soon as $\mu^2 > 4(q-1)$ we have $x^\star(\mu) \in \mathbb{R}$, hence for all such eigenvalues $\mathrm{Arg}\big(x^\star(\mu) + a \big) = 0$ for any real a, and thus $\varrho_0(\mu) = 0$. For eigenvalues $\mu^2 < 4(q-1)$, on the other hand, we may use identities such as $\mathrm{dArg}(z)/\mathrm{d}\mu = \mathrm{Im}(z^{-1}\mathrm{d}z/\mathrm{d}\mu)$ and $\mathrm{d}x^\star(\mu)/\mathrm{d}\mu =$

$-\mathrm{i}x^\star(\mu)/\sqrt{4(q-1)-\mu^2}$ to derive:

$$\begin{aligned}\varrho_0(\mu) &= \frac{1}{2\pi}\mathrm{Im}\Big\{\frac{\mathrm{d}x^\star(\mu)}{\mathrm{d}\mu}\Big[\frac{q-2}{x^\star(\mu)+1}+\frac{q-2}{x^\star(\mu)-1}-\frac{2(q-1)}{x^\star(\mu)}\Big]\Big\}\\ &= \frac{1}{\pi\sqrt{4(q-1)-\mu^2}}\,\mathrm{Re}\Big\{q-1-(q-2)\frac{x^{\star 2}(\mu)}{x^{\star 2}(\mu)-1}\Big\}\\ &= \frac{q}{2\pi}\frac{\sqrt{4(q-1)-\mu^2}}{q^2-\mu^2}. \end{aligned} \tag{C.41}$$

Hence, in combination,

$$\varrho_0(\mu) = \frac{q}{2\pi}\frac{\sqrt{4(q-1)-\mu^2}}{q^2-\mu^2}\,\theta\big[2\sqrt{q-1}-|\mu|\big]. \tag{C.42}$$

In the same way we derive expression (5.52) for the function $h(\mu)$:

$$h(\mu) = -\frac{1}{\pi}\mathrm{Im}\Big[\frac{-\mathrm{i}x^\star(\mu)}{\sqrt{4(q-1)-\mu^2}}\Big] = \frac{1}{\pi}\frac{\theta\left(2\sqrt{q-1}-|\mu|\right)}{\sqrt{4(q-1)-\mu^2}}. \tag{C.43}$$

C.4 Expected number of subgraphs in a RRG

Here we restate and apply a result in [71] on the expected number $\mathrm{E}[\mathbf{J}]$ of strictly balanced subgraphs $\mathbf{J}$ with k nodes and ℓ edges, in a random regular graph $\mathbf{A}$ with N nodes and degree q (see [71] for the precise definition of strictly balanced subgraphs, here we only require that these include loops and cliques). This number is given by

$$\mathrm{E}[\mathbf{J}] = P(\mathbf{J}\subset\mathbf{A})\frac{[N]_k}{a(\mathbf{J})}, \qquad P(\mathbf{J}\subset\mathbf{A}) = \frac{\prod_{i=1}^N [q]_{j_i}}{(qN)^\ell}\Big(1+\mathcal{O}\big((\frac{qk}{N})^2\big)\Big) \tag{C.44}$$

Here $P(\mathbf{J}\subset\mathbf{A})$ is the probability of $\mathbf{J}$ being a subgraph of $\mathbf{A}$, j_i is the degree of node i when computed only via incident links that belong to $\mathbf{J}$, $[r]_s = r!/(r-s)!$, and $a(\mathbf{J})$ is the number of automorphisms of $\mathbf{J}$. For the case of a length-ℓ loop, $\mathbf{J}=\mathbf{A}_\ell$, we have $k=\ell$ and $a(\mathbf{J})=2\ell$, and hence

$$\mathrm{E}[\mathbf{A}_\ell] = (q-1)^\ell/2\ell \tag{C.45}$$

For $(q+1)$-node cliques we have $k=q+1$, $\ell=\frac{1}{2}q(q+1)$, and $a(\mathbf{J})=(q+1)!$, so

$$\mathrm{E}[K_{q+1}] = \frac{(q!)^{q+1}}{(Nq)^{\frac{q}{2}(q+1)}}\frac{[N]_{q+1}}{(q+1)!} \tag{C.46}$$

Appendix D

Functional integral for arbitrary degree distribution $p(k)$

In this Appendix we simplify expression (6.5,6.6) for our generating function $\phi[\hat{\varrho}]$.

The first part has a very strong resemblance to Appendix C.1, nevertheless the derivation is not exactly the same. We have included a different version of the calculation to show how different choices lead to the same result. Additionally some definitions are slightly modified to better suit the general case.

D.1 Tranformation to Fourier components

We first introduce a number of definitions to compactify our formulae. We begin by introducing the Fourier coefficients of our functional order parameters:

$$P(\boldsymbol{\varphi},\omega)=\sum_{\ell\in\mathbb{Z}}W_\ell(\boldsymbol{\varphi})\frac{\mathrm{e}^{\mathrm{i}\ell\omega}}{2\pi},\qquad W_\ell(\boldsymbol{\varphi})=\int_{-\pi}^{\pi}\mathrm{d}\omega\ P(\boldsymbol{\varphi},\omega)\mathrm{e}^{-\mathrm{i}\ell\omega},\tag{D.1}$$

$$\hat{P}(\boldsymbol{\varphi},\omega)=\sum_{\ell\in\mathbb{Z}}\hat{W}_\ell(\boldsymbol{\varphi})\mathrm{e}^{-\mathrm{i}\ell\omega},\qquad \hat{W}_\ell(\boldsymbol{\varphi})=\int_{-\pi}^{\pi}\frac{\mathrm{d}\omega}{2\pi}\ \hat{P}(\boldsymbol{\varphi},\omega)\mathrm{e}^{\mathrm{i}\ell\omega}.\tag{D.2}$$

We then write $S[P,\hat{P}]$ strictly in terms of the Fourier transforms $\{W_\ell(\boldsymbol{\varphi}),\hat{W}_\ell(\boldsymbol{\varphi})\}$, noting that $\omega\in[-\pi,\pi]$. The result is

$$S[P,\hat{P}]=S[W_1,\hat{W}_1,W_2,\hat{W}_2]+S_0[\{W_\ell,\hat{W}_\ell\}],\tag{D.3}$$

where, using the notation $\mathbf{r} = \{r_\ell \geq 0,\ \ell \in \mathbb{Z}\}$,

$$\begin{aligned} S_0[\{W_\ell, \hat{W}_\ell\}] &= \mathrm{i} \sum_{\ell \notin \{1,2\}} \int \mathrm{d}\boldsymbol{\varphi}\ W_\ell(\boldsymbol{\varphi}) \hat{W}_\ell(\boldsymbol{\varphi}) \\ &\quad + \sum_k p(k) \log \int \mathrm{d}\boldsymbol{\varphi}\ \nu(\boldsymbol{\varphi}) \sum_{\mathbf{r}} \delta_{k, \sum_{\ell \in \mathbb{Z}} \ell r_\ell} \prod_{\ell \in \mathbb{Z}} \frac{[-\mathrm{i}\hat{W}_\ell(\boldsymbol{\varphi})]^{r_\ell}}{r_\ell!}, \quad \text{(D.4)} \\ S[W_1, \hat{W}_1, W_2, \hat{W}_2] &= -\frac{c}{2}(1+\frac{c-2}{2N}) + \mathrm{i} \int \mathrm{d}\boldsymbol{\varphi} \Big[W_1(\boldsymbol{\varphi}) \hat{W}_1(\boldsymbol{\varphi}) + W_2(\boldsymbol{\varphi}) \hat{W}_2(\boldsymbol{\varphi}) \Big] \\ &\quad - \frac{c}{2N} \int \mathrm{d}\boldsymbol{\varphi}\ W_2(\boldsymbol{\varphi}) \mathrm{e}^{-\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}} + \frac{c}{2}(1+\frac{c}{N}) \int \mathrm{d}\boldsymbol{\varphi} \mathrm{d}\boldsymbol{\varphi}'\ W_1(\boldsymbol{\varphi}) W_1(\boldsymbol{\varphi}') \mathrm{e}^{-\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'} \\ &\quad - \frac{c^2}{4N} \int \mathrm{d}\boldsymbol{\varphi} \mathrm{d}\boldsymbol{\varphi}'\ W_2(\boldsymbol{\varphi}) W_2(\boldsymbol{\varphi}') \mathrm{e}^{-2\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'}. \quad \text{(D.5)} \end{aligned}$$

The integration in (6.5) can be replaced by integration over the functional Fourier components, in exactly the same manner as in Apppendix C.1.

Then the generating function can be written as, modulo an irrelevant additive constant:

$$\phi[\hat{\varrho}] = \lim \frac{1}{N} \log \int \Big[\prod_{\ell \in \mathbb{Z}} \mathcal{D}W_\ell \mathcal{D}\hat{W}_\ell \Big] \mathrm{e}^{NS[W_1, \hat{W}_1, W_2, \hat{W}_2] + NS_0[\{W_\ell, \hat{W}_\ell\}]}. \quad \text{(D.6)}$$

With (A.6), integration over the Fourier components with $\ell = 1, 2$ has become trivial:

$$\begin{aligned} &\int \prod_{\ell \notin \{1,2\}} [\mathcal{D}W_\ell \mathcal{D}\hat{W}_\ell]\ \mathrm{e}^{NS_0[\{W_\ell, \hat{W}_\ell\}]} = \mathrm{e}^{\mathrm{i}N \sum_{\ell=1}^2 \int \mathrm{d}\boldsymbol{\varphi}\ W_\ell(\boldsymbol{\varphi}) \hat{W}_\ell(\boldsymbol{\varphi})} \\ &\times \int \prod_{\ell \notin \{1,2\}} \big[\mathcal{D}\hat{W}_\ell\ \delta[\hat{W}_\ell]\big] \exp\Big(N \sum_k p(k) \log \int \mathrm{d}\boldsymbol{\varphi}\ \nu(\boldsymbol{\varphi}) \sum_{\mathbf{r}} \delta_{k, \sum_{\ell \in \mathbb{Z}} \ell r_\ell} \prod_{\ell \in \mathbb{Z}} \frac{[-\mathrm{i}\hat{W}_\ell(\boldsymbol{\varphi})]^{r_\ell}}{r_\ell!} \Big) \\ &= \exp\Big(\mathrm{i}N \sum_{\ell=1}^2 \int \mathrm{d}\boldsymbol{\varphi}\ W_\ell(\boldsymbol{\varphi}) \hat{W}_\ell(\boldsymbol{\varphi}) \\ &+ N \sum_k p(k) \log \int \mathrm{d}\boldsymbol{\varphi}\ \nu(\boldsymbol{\varphi}) \sum_{0 \leq r \leq k/2} \frac{[-\mathrm{i}\hat{W}_1(\boldsymbol{\varphi})]^{k-2r}}{(k-2r)!} \frac{[-\mathrm{i}\hat{W}_2(\boldsymbol{\varphi})]^r}{r!} \Big) \Big]. \quad \text{(D.7)} \end{aligned}$$

From experience we know that $\hat{W}_2$ will be associated with the $\mathcal{O}(1/N)$ fluctuations, so we introduce the scaling $\hat{W}_2 \to \hat{W}/N$. This allows us to discard the $r \geq 2$ in the sum over r and to expand the logarithm. Hence (6.5) can be written as follows:

$$\begin{aligned}\phi[\hat{\varrho}] =&\frac{1}{N}\log\int \mathcal{D}W_1\mathcal{D}W_2\mathcal{D}\hat{W}_1\mathcal{D}\hat{W}_2\,\mathrm{e}^{-\frac{Nc}{2}-\frac{c-2}{2}+N\mathrm{i}\int \mathrm{d}\boldsymbol{\varphi}\, W_1(\boldsymbol{\varphi})\hat{W}_1(\boldsymbol{\varphi})+\mathrm{i}\int \mathrm{d}\boldsymbol{\varphi}\, W_2(\boldsymbol{\varphi})\hat{W}_2(\boldsymbol{\varphi})}\\ &\times \mathrm{e}^{\frac{Nc}{2}\left(1+\frac{c}{N}\right)\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}' W_1(\boldsymbol{\varphi})U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W_1(\boldsymbol{\varphi}')-\frac{c}{2}\int \mathrm{d}\boldsymbol{\varphi} W_2(\boldsymbol{\varphi})V(\boldsymbol{\varphi})-\frac{c^2}{4}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}' W_2(\boldsymbol{\varphi})U_2(\boldsymbol{\varphi},\boldsymbol{\varphi}')W_2(\boldsymbol{\varphi}')}\\ &\times \mathrm{e}^{\mathrm{i}\int \mathrm{d}\boldsymbol{\varphi}\mathcal{R}[\hat{W}_1(\boldsymbol{\varphi})]\hat{W}_2(\boldsymbol{\varphi})+N\sum_k p(k)\log\int \mathrm{d}\boldsymbol{\varphi}\,\nu(\boldsymbol{\varphi})\frac{1}{k!}[-\mathrm{i}\hat{W}_1(\boldsymbol{\varphi})]^k}\\ =&\frac{1}{N}\log\int \mathcal{D}\hat{W}_1\mathrm{e}^{-\frac{Nc}{2}-\frac{c-2}{2}+N\sum_k p(k)\log\int \mathrm{d}\boldsymbol{\varphi}\,\nu(\boldsymbol{\varphi})\frac{1}{k!}[-\mathrm{i}\hat{W}_1(\boldsymbol{\varphi})]^k}\\ &\times\int \mathcal{D}W_1\,\mathrm{e}^{\frac{Nc}{2}\left(1+\frac{c}{N}\right)\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}' W_1(\boldsymbol{\varphi})U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W_1(\boldsymbol{\varphi}')+N\mathrm{i}\int \mathrm{d}\boldsymbol{\varphi} W_1(\boldsymbol{\varphi})\hat{W}_1(\boldsymbol{\varphi})}\\ &\times\int \mathcal{D}W_2\,\mathrm{e}^{-\frac{c^2}{4}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}' W_2(\boldsymbol{\varphi})U_2(\boldsymbol{\varphi},\boldsymbol{\varphi}')W_2(\boldsymbol{\varphi}')-\frac{c}{2}\int \mathrm{d}\boldsymbol{\varphi} W_2(\boldsymbol{\varphi})V(\boldsymbol{\varphi})}\\ &\times\int \mathcal{D}\hat{W}_2\,\mathrm{e}^{\mathrm{i}\int \mathrm{d}\boldsymbol{\varphi}\left[\mathcal{R}[\hat{W}_1(\boldsymbol{\varphi})]+W_2(\boldsymbol{\varphi})\right]\hat{W}_2(\boldsymbol{\varphi})}.\end{aligned}\tag{D.8}$$

Where we have defined,

$$\mathcal{R}[\hat{W}_1(\boldsymbol{\varphi})]=\sum_k p(k)k(k-1)\frac{\nu(\boldsymbol{\varphi})(\hat{W}_1(\boldsymbol{\varphi}))^{k-2}}{\int \mathrm{d}\boldsymbol{\varphi}'\,\nu(\boldsymbol{\varphi}')(\hat{W}_1(\boldsymbol{\varphi}'))^k}.\tag{D.9}$$

The integral over $\hat{W}_2$ gives a functional delta, $\delta[\mathcal{R}[\hat{W}_1(\boldsymbol{\varphi})]+W_2(\boldsymbol{\varphi})]$, so the integral over W_2 becomes trivial. It is just a matter of substituting $W_2(\boldsymbol{\varphi})$ with $-\mathcal{R}[\hat{W}_1(\boldsymbol{\varphi})]$. The integral over W_1 is a simple Gaussian integral.

$$\begin{aligned}\phi[\hat{\varrho}] =&\frac{1}{N}\log\sqrt{\det\left([c\mathbf{U}_1]^{-1}\right)}\int \mathcal{D}\hat{W}_1\,\mathrm{e}^{-\frac{Nc}{2}-\frac{c-2}{2}+N\sum_k p(k)\log\int \mathrm{d}\boldsymbol{\varphi}\nu(\boldsymbol{\varphi})\frac{1}{k!}[-\mathrm{i}\hat{W}_1(\boldsymbol{\varphi})]^k}\\ &\times \mathrm{e}^{\frac{N}{2c}\left(1-\frac{c}{N}\right)\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\,\hat{W}_1(\boldsymbol{\varphi}){U_1}^{-1}(\boldsymbol{\varphi},\boldsymbol{\varphi}')\hat{W}_1(\boldsymbol{\varphi}')-\frac{c^2}{4}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\mathcal{R}[\hat{W}_1(\boldsymbol{\varphi})]U_2(\boldsymbol{\varphi},\boldsymbol{\varphi}')\mathcal{R}[\hat{W}_1(\boldsymbol{\varphi}')]}\\ &\times \mathrm{e}^{\frac{c}{2}\int \mathrm{d}\boldsymbol{\varphi}\,\mathcal{R}[\hat{W}_1(\boldsymbol{\varphi})]V(\boldsymbol{\varphi})}.\end{aligned}\tag{D.10}$$

We now introduce the following change of variable,

$$\hat{W}_1(\boldsymbol{\varphi})=\mathrm{i}c\int \mathrm{d}\boldsymbol{\varphi}' U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}').\tag{D.11}$$

Note that the Jacobian is precisely the determinant outside the integral,

$$\left|\frac{\partial\hat{W}_1}{\partial W}\right|\propto\left|\det\left(c\mathbf{U}_1\right)\right|.\tag{D.12}$$

Therefore, getting rid of constants, we can rewrite our integral in the following way.

$$\begin{aligned}
\phi[\hat{\varrho}] =& \frac{1}{N}\log\sqrt{\det\left(c\mathbf{U}_1\right)}\int \mathcal{D}W\,\mathrm{e}^{NS_0[W]+S_1[W]}\\
S_0[W] =& \sum_k p(k)\log\int \mathrm{d}\boldsymbol{\varphi}\,\nu(\boldsymbol{\varphi})\frac{1}{k!}\left(\int \mathrm{d}\boldsymbol{\varphi}' U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}')\right)^k\\
&-\frac{c}{2}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\,W(\boldsymbol{\varphi})U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}')\\
S_1[W] =& \frac{c^2}{2}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\,W(\boldsymbol{\varphi})U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}')-\frac{1}{4}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}'\,r(\boldsymbol{\varphi})U_2(\boldsymbol{\varphi},\boldsymbol{\varphi}')r(\boldsymbol{\varphi}')\\
&-\frac{1}{2}\int \mathrm{d}\boldsymbol{\varphi}\,r(\boldsymbol{\varphi})U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}).
\end{aligned}\tag{D.13}$$

Here

$$\begin{aligned}
r(\boldsymbol{\varphi}) =& -c\mathcal{R}\left[\mathrm{i}c\int \mathrm{d}\boldsymbol{\varphi}' U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}')\right]\\
=& \sum_k p(k)\frac{k}{c}(k-1)\frac{\nu(\boldsymbol{\varphi})[\int \mathrm{d}\boldsymbol{\varphi}' U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}')]^{k-2}}{\int \mathrm{d}\boldsymbol{\varphi}''\,\nu(\boldsymbol{\varphi}'')[\int \mathrm{d}\boldsymbol{\varphi}'\,U_1(\boldsymbol{\varphi}'',\boldsymbol{\varphi}')W(\boldsymbol{\varphi}')]^k}.
\end{aligned}\tag{D.14}$$

D.2 Saddle point approximation

It is convenient to calculate the derivatives of the leading order term. We introduce the notation $Z_k = \int \mathrm{d}\boldsymbol{\varphi}\nu(\boldsymbol{\varphi})[\int \mathrm{d}\boldsymbol{\varphi}' U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}')]^k$ for convenience.

$$\begin{aligned}
\frac{\delta S_0}{\delta W(\boldsymbol{\varphi})} =& \sum_k p(k)k\frac{1}{Z_k}\int \mathrm{d}\boldsymbol{\varphi}' U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')\nu(\boldsymbol{\varphi}')\left[\int \mathrm{d}\boldsymbol{\varphi}''\,U_1(\boldsymbol{\varphi}',\boldsymbol{\varphi}'')W(\boldsymbol{\varphi}'')\right]^{k-1}\\
&-c\int \mathrm{d}\boldsymbol{\varphi}U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W(\boldsymbol{\varphi}'),\\
\frac{\delta^2 S}{\delta W(\boldsymbol{\varphi})\delta W(\boldsymbol{\psi})} =& \sum_k p(k)k(k-1)\frac{1}{Z_k}\int \mathrm{d}\boldsymbol{\varphi}'\,U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')U_1(\boldsymbol{\varphi}',\boldsymbol{\psi})\nu(\boldsymbol{\varphi}')\\
&\qquad\times\left[\int \mathrm{d}\boldsymbol{\varphi}''\,U_1(\boldsymbol{\varphi}',\boldsymbol{\varphi}'')W(\boldsymbol{\varphi}'')\right]^{k-2}\\
&-\sum_k p(k)k^2\frac{1}{{Z_k}^2}\int \mathrm{d}\boldsymbol{\varphi}' U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')\nu(\boldsymbol{\varphi}')\left[\int \mathrm{d}\boldsymbol{\varphi}''\,U_1(\boldsymbol{\varphi}',\boldsymbol{\varphi}'')W(\boldsymbol{\varphi}'')\right]^{k-1}\\
&\qquad\times\int \mathrm{d}\boldsymbol{\varphi}' U_1(\boldsymbol{\psi},\boldsymbol{\varphi}')\nu(\boldsymbol{\varphi}')\left[\int \mathrm{d}\boldsymbol{\varphi}''\,U_1(\boldsymbol{\varphi}',\boldsymbol{\varphi}'')W(\boldsymbol{\varphi}'')\right]^{k-1}\\
&-cU_1(\boldsymbol{\varphi},\boldsymbol{\psi})
\end{aligned}\tag{D.15}$$

We can perform the integral by saddle point approximation. If we assume fluctuations

around the stationary point W_0 are $\mathcal{O}\left(1/\sqrt{N}\right)$ we can do the following substitution, $W = W_0 + \frac{1}{\sqrt{N}}W_1$ and include the Gaussian integral,

$$\phi[\hat{\varrho}] = \frac{1}{N}\log\sqrt{\det(c\mathbf{U}_1)}\int \mathcal{D}W \mathrm{e}^{NS_0[W]+S_1[W]} \sim S_0[W_0] + \frac{1}{N}S_1[W_0] + \frac{1}{N}\log\sqrt{\frac{\det(c\mathbf{U}_1)}{\det(-\mathbf{\Gamma})}} + \mathcal{O}\left(1/N^{\frac{3}{2}}\right) \tag{D.16}$$

Denoting $\mathbf{\Gamma} = \delta^2 S$ the Hessian of S evaluated at W_0.

D.2.1 Saddle point equation

The saddle point needs to be calculated with the leading order term, that is

$$\frac{\delta S_0}{\delta W}[W_0] = 0. \tag{D.17}$$

This leads to the following equations,

$$\begin{aligned} W_0(\boldsymbol{\varphi}) =& \sum_k p(k)\frac{k}{c}\frac{F_{k-1}(\boldsymbol{\varphi})}{Z_k} \\ F_k(\boldsymbol{\varphi}) =& \nu(\boldsymbol{\varphi})\left[\int \mathrm{d}\boldsymbol{\varphi}' U_1(\boldsymbol{\varphi},\boldsymbol{\varphi}')W_0(\boldsymbol{\varphi}')\right]^k. \end{aligned} \tag{D.18}$$

D.2.2 Hessian, the $\mathbf{\Gamma}$ operator

Evaluated at the stationary solution, W_0, we can rewrite $\mathbf{\Gamma}$ in the following way.

$$\begin{aligned} \mathbf{\Gamma} =& -c\mathbf{U}_1 + c\mathbf{U}_1\mathbf{T} \\ \mathbf{T} =& \mathbf{M} - \mathbf{B} \\ M(\boldsymbol{\varphi},\boldsymbol{\psi}) =& r_s(\boldsymbol{\varphi})U_1(\boldsymbol{\varphi},\boldsymbol{\psi}) \\ B(\boldsymbol{\varphi},\boldsymbol{\psi}) =& \sum_k p(k)\frac{k^2}{c}\frac{1}{{Z_k}^2}F_{k-1}(\boldsymbol{\varphi})\cdot\int \mathrm{d}\boldsymbol{\varphi}' U_1(\boldsymbol{\psi},\boldsymbol{\varphi}')F_{k-1}(\boldsymbol{\varphi}') \\ r_s(\boldsymbol{\varphi}) =& \sum_k p(k)\frac{k}{c}(k-1)\frac{F_{k-2}(\boldsymbol{\varphi})}{\cdot} \end{aligned} \tag{D.19}$$

Finally,

$$\begin{aligned}\phi[\hat{\varrho}] =& \lim \Bigg[\sum_k p(k) \log Z_k - \frac{c}{2} \int \mathrm{d}\boldsymbol{\varphi} \mathrm{d}\boldsymbol{\varphi}' \, W_0(\boldsymbol{\varphi}) U_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}') W_0(\boldsymbol{\varphi}') \\ &+ \frac{1}{N} \left(\frac{c^2}{2} \int \mathrm{d}\boldsymbol{\varphi} \mathrm{d}\boldsymbol{\varphi}' W_0(\boldsymbol{\varphi}') U_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}') W(\boldsymbol{\varphi}) - \frac{1}{4} \int \mathrm{d}\boldsymbol{\varphi} \mathrm{d}\boldsymbol{\varphi}' \, r_s(\boldsymbol{\varphi}) U_2(\boldsymbol{\varphi}, \boldsymbol{\varphi}') r_s(\boldsymbol{\varphi}') \right. \\ &\left. - \frac{1}{2} \int \mathrm{d}\boldsymbol{\varphi} \, r_s(\boldsymbol{\varphi}) U_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}) + \sum_{\ell=1}^{\infty} \frac{1}{2\ell} \mathrm{Tr}((\mathbf{M} - \mathbf{B})^{\ell}) \right) \Bigg]. \end{aligned} \tag{D.20}$$

This expression will show its full meaning once the RS assumption is made.

D.3 Replica symmetric theory

We introduce the following replica symmetric (RS) ansatz.

$$\begin{aligned} W_0(\boldsymbol{\varphi}) =& \mathcal{C} \int \mathrm{d}\mathbf{X} \, W(\mathbf{X}) \frac{1}{Z(\mathbf{X})} \mathrm{e}^{-\frac{\mathrm{i}}{2} \boldsymbol{\varphi} \cdot \mathbf{X} \boldsymbol{\varphi}} \\ Z(\mathbf{X}) =& \int \mathrm{d}\boldsymbol{\varphi} \, \mathrm{e}^{-\frac{\mathrm{i}}{2} \boldsymbol{\varphi} \cdot \mathbf{X} \boldsymbol{\varphi}} = \prod_{\mu} \left(\frac{2\pi}{\mathrm{i} x(\mu)} \right)^{\frac{n_\mu}{2}} \left(\overline{\frac{2\pi}{\mathrm{i} x(\mu)}} \right)^{\frac{m_\mu}{2}}. \end{aligned} \tag{D.21}$$

Substituting into the saddle point equation leads to the following equations.

$$\begin{aligned} W(\mathbf{X}) =& \frac{1}{\mathcal{C}^2} \sum_k p(k) \frac{k}{c} \frac{Z(\mathbf{X}) f_{k-1}(\mathbf{X})}{\mathcal{Z}_k} \\ \mathcal{C}^2 =& \sum_k p(k) \frac{k}{c} \frac{\mathcal{Z}_{k-1}}{\mathcal{Z}_k} \\ f_k(\mathbf{X}) =& \int \prod_{\ell=1}^{k} \mathrm{d}\mathbf{X}_\ell W(\mathbf{X}_\ell) \cdot \delta(\mathbf{X} + \boldsymbol{\mu} + \sum_{\ell=1}^{k} \mathbf{X}_\ell^{-1}) \\ \mathcal{Z}_k =& \int \mathrm{d}\mathbf{X} \, f_k(\mathbf{X}) Z(\mathbf{X}) = \int \prod_{\ell=1}^{k} \mathrm{d}\mathbf{X}_\ell W(\mathbf{X}_\ell) \cdot Z(-\boldsymbol{\mu} - \sum_{\ell=1}^{k} \mathbf{X}_\ell^{-1}). \end{aligned} \tag{D.22}$$

These equations lead to the following saddle point action, the leading order.

$$S_0[W] = c \log \mathcal{C} + \sum_k p(k) \log \mathcal{Z}_k - \frac{c}{2} \mathcal{C}^2 \int \mathrm{d}\mathbf{X} \mathrm{d}\mathbf{X}' \, W(\mathbf{X}) W(\mathbf{X}') \mathcal{K}(\mathbf{X}, \mathbf{X}'). \tag{D.23}$$

Where

$$\mathcal{K}(\mathbf{X},\mathbf{X}') = \frac{Z(\mathbf{X}-\mathbf{X}'^{-1})}{Z(\mathbf{X})} = \prod_{\mu}\left(\frac{1}{1-\frac{1}{x(\mu)x'(\mu)}}\right)^{\frac{n_\mu}{2}}\left(\overline{\frac{1}{1-\frac{1}{x(\mu)x'(\mu)}}}\right)^{\frac{m_\mu}{2}} = \frac{Z(\mathbf{X}'-\mathbf{X}^{-1})}{Z(\mathbf{X}')}. \tag{D.24}$$

From W satisfying the saddle point equations we can prove the following identity,

$$\mathcal{C}^{-2} = \int \mathrm{d}\mathbf{X}\mathrm{d}\mathbf{X}'\, W(\mathbf{X})W(\mathbf{X}')\mathcal{K}(\mathbf{X},\mathbf{X}'). \tag{D.25}$$

Note it is equivalent to proving

$$\mathcal{C}^{2}\int \mathrm{d}\mathbf{X}\mathrm{d}\mathbf{X}'\, W(\mathbf{X})W(\mathbf{X}')\mathcal{K}(\mathbf{X},\mathbf{X}') = 1. \tag{D.26}$$

Which can be done by simply using the saddle point equation.

$$\begin{aligned}
&\mathcal{C}^{2}\int \mathrm{d}\mathbf{X}\mathrm{d}\mathbf{X}'\, W(\mathbf{X})W(\mathbf{X}')\mathcal{K}(\mathbf{X},\mathbf{X}') = \\
&= \int \mathrm{d}\mathbf{X}\mathrm{d}\mathbf{X}' W(\mathbf{X})\sum_k p(k)\frac{k}{c}\frac{Z(\mathbf{X}')f_{k-1}(\mathbf{X}')}{\mathcal{Z}_k}\frac{Z(\mathbf{X}'-\mathbf{X}^{-1})}{Z(\mathbf{X}')} = \\
&= \sum_k p(k)\frac{k}{c}\frac{1}{\mathcal{Z}_k}\int \mathrm{d}\mathbf{X}'\int \mathrm{d}\mathbf{X} W(\mathbf{X})f_{k-1}(\mathbf{X}')Z(\mathbf{X}'-\mathbf{X}^{-1}) = \sum_k p(k)\frac{k}{c}\frac{\mathcal{Z}_k}{\mathcal{Z}_k} = 1.
\end{aligned} \tag{D.27}$$

Therefore,

$$S_0[W] = -\frac{c}{2}\log\int \mathrm{d}\mathbf{X}\mathrm{d}\mathbf{X}'\, W(\mathbf{X})W(\mathbf{X}')\mathcal{K}(\mathbf{X},\mathbf{X}') + \sum_k p(k)\log\mathcal{Z}_k. \tag{D.28}$$

For the subleading terms we first need to calculate the RS forms of $r(\boldsymbol{\varphi})$ and $B(\boldsymbol{\varphi},\boldsymbol{\psi})$, which are given by

$$\begin{aligned}
r(\boldsymbol{\varphi}) =& \frac{1}{\mathcal{C}^2}\sum_k p(k)\frac{k}{c}(k-1)\frac{1}{\mathcal{Z}_k}\int \mathrm{d}\mathbf{X} f_{k-2}(\mathbf{X})\mathrm{e}^{-\frac{\mathrm{i}}{2}\boldsymbol{\varphi}\cdot\mathbf{X}\boldsymbol{\varphi}},\\
B(\boldsymbol{\varphi},\boldsymbol{\psi}) =& \frac{1}{\mathcal{C}^2}\sum_k p(k)\frac{k^2}{c}\frac{1}{{\mathcal{Z}_k}^2}\int \mathrm{d}\boldsymbol{\varphi}'\int \mathrm{d}\mathbf{X}\mathrm{d}\mathbf{X}' f_{k-1}(\mathbf{X})f_{k-1}(\mathbf{X}')\mathrm{e}^{-\frac{\mathrm{i}}{2}\boldsymbol{\varphi}\cdot\mathbf{X}\boldsymbol{\varphi}-\mathrm{i}\boldsymbol{\psi}\cdot\boldsymbol{\varphi}'-\frac{\mathrm{i}}{2}\boldsymbol{\varphi}'\cdot\mathbf{X}'\boldsymbol{\varphi}'}.
\end{aligned} \tag{D.29}$$

Once we have calculated these, we can then proceed to calculate the traces of the products of operators $\mathbf{M}$ and $\mathbf{B}$.

Let us look at the pure traces of $\mathbf{M}$ first,

$$\begin{aligned}
\mathrm{Tr}(\mathbf{M}^\ell) &= \int \prod_{p=1}^{\ell} \mathrm{d}\boldsymbol{\varphi}_p \prod_{p=1}^{\ell} r(\boldsymbol{\varphi}_p) U_1(\boldsymbol{\varphi}_p, \boldsymbol{\varphi}_{p+1}) \\
&= \left[\prod_{p=1}^{\ell} \sum_{k_p}\right] \left[\prod_{p=1}^{\ell} \tilde{p}(k_p)(k_p-1)\right] \prod_{p=1}^{\ell} \frac{1}{\mathcal{Z}_{k_p-1}} \\
&\qquad \times \int \prod_{p=1}^{\ell} \mathrm{d}\mathbf{X}_p f_{k_p-2}(\mathbf{X}_p) \mathrm{d}\boldsymbol{\varphi}_p \, \mathrm{e}^{-\frac{\mathrm{i}}{2}\sum_p \boldsymbol{\varphi}_p \cdot \mathbf{X}_p \boldsymbol{\varphi}_p - \mathrm{i}\sum_p \boldsymbol{\varphi}_p \cdot \boldsymbol{\varphi}_{p+1}} \\
&= \left[\prod_{p=1}^{\ell} \sum_{k_p}\right] \left[\prod_{p=1}^{\ell} \tilde{p}(k_p)(k_p-1)\right] \prod_{p=1}^{\ell} \frac{1}{\mathcal{Z}_{k_p-1}} \\
&\qquad \times \int \prod_{p=1}^{\ell} \mathrm{d}\mathbf{X}_p f_{k_p-2}(\mathbf{X}_p) \prod_{\mu} \prod_{\alpha_\mu=1}^{n_\mu} \left(\int \mathrm{d}^\ell \boldsymbol{\phi}_{\mu,\alpha_\mu} \, \mathrm{e}^{-\frac{\mathrm{i}}{2} \boldsymbol{\phi}_{\mu,\alpha_\mu} \cdot (\mathbf{A}^*_{\ell,\mu} - \mu_\varepsilon \mathbf{I}_\ell) \boldsymbol{\phi}_{\mu,\alpha_\mu}} \right) \\
&\qquad\qquad \times \prod_{\beta_\mu=1}^{m_\mu} \overline{\left(\int \mathrm{d}^\ell \boldsymbol{\psi}_{\mu,\beta_\mu} \, \mathrm{e}^{-\frac{\mathrm{i}}{2} \boldsymbol{\psi}_{\mu,\beta_\mu} \cdot (\mathbf{A}^*_{\ell,\mu} - \mu_\varepsilon \mathbf{I}_\ell) \boldsymbol{\psi}_{\mu,\beta_\mu}} \right)} \\
&= \left[\prod_{p=1}^{\ell} \sum_{k_p} \tilde{p}(k_p)(k_p-1)\right] \mathrm{e}^{\ell R(\{k_p\}_{p=1,\ldots,\ell})}.
\end{aligned} \tag{D.30}$$

Where we have defined

$$(\mathbf{A}^*_{\ell,\mu})_{p,p'} = \delta_{p,p'-1} + \delta_{p,p'+1} + (\mu_\varepsilon + x_p(\mu))\delta_{p,p} \quad (\text{with } p \bmod \ell). \tag{D.31}$$

This expression corresponds to a loop of length ℓ with complex fields of value $\mu + x_p(\mu)$ in the diagonal, where $x_p(\mu)$ is distributed according to $f_{k_p-2}[\boldsymbol{x}_p]$. Due to the definition of $f_{k_p-2}[\boldsymbol{x}]$, this means that each node is receiving $k_p - 2$ "messages" of value $-1/x(\mu)$ where each $\boldsymbol{x}$ is distributed according to $W[\boldsymbol{x}]$. Where,

$$\begin{aligned}
R(\{k_p\}_{p=1,\ldots,\ell}) =& \frac{1}{\ell} \log \int \prod_{p=1}^{\ell} \mathrm{d}\mathbf{X}_p f_{k_p-2}(\mathbf{X}_p) \\
&\qquad \times \prod_{\mu} \left(Z(\mu_\varepsilon | \mathbf{A}^*_{\ell,\mu})\right)^{n_\mu} \overline{\left(Z(\mu_\varepsilon | \mathbf{A}^*_{\ell,\mu})\right)}^{m_\mu} \\
&- \frac{1}{\ell} \sum_{p=1}^{\ell} \log \mathcal{Z}_{k_p-1}.
\end{aligned} \tag{D.32}$$

In order to understand how to write down the other elements, we will proceed by writing out a few simpler examples before the general expression.

We can start with the following case of just one $\mathbf{B}$ and the rest all $\mathbf{M}$'s,

$$\begin{aligned}\mathrm{Tr}(\mathbf{M}^{\ell-1}\mathbf{B}) &= \int \prod_{p=1}^{\ell} \mathrm{d}\boldsymbol{\varphi}_p \cdot \left[\prod_{p=1}^{\ell}\sum_{k_p} \tilde{p}(k_p)(k_p-1+\delta_{p,\ell})\right] \frac{1}{\mathcal{Z}_{k_\ell}} \prod_{p=1}^{\ell} \frac{1}{\mathcal{Z}_{k_p-1}} \\ &\times \int \mathrm{d}\boldsymbol{\psi}\mathrm{d}\mathbf{X} f_{k_\ell-1}(\mathbf{X}) \prod_{p=1}^{\ell} \mathrm{d}\mathbf{X}_p f_{k_p-2+\delta_{p,\ell}}(\mathbf{X}_p)\, \mathrm{e}^{-\frac{\mathrm{i}}{2}\sum_{p=1}^{\ell}\boldsymbol{\varphi}_p\cdot\mathbf{X}_p\boldsymbol{\varphi}_p - \mathrm{i}\sum_{p=1}^{\ell-1}\boldsymbol{\varphi}_p\cdot\boldsymbol{\varphi}_{p+1} - \mathrm{i}\boldsymbol{\varphi}_1\cdot\boldsymbol{\psi} - \frac{\mathrm{i}}{2}\boldsymbol{\psi}\cdot\mathbf{X}\boldsymbol{\psi}} \\ &= \left[\prod_{p=1}^{\ell}\sum_{k_p} \tilde{p}(k_p)(k_p-1+\delta_{p,\ell})\right] \mathrm{e}^{(\ell+1)C(k_\ell,\{k_p\}_{p=1,\ldots,\ell})}. \end{aligned} \tag{D.33}$$

Where,

$$\begin{aligned} C(k,\{k_p\}_{p=1,\ldots,\ell}) =& \frac{1}{\ell+1}\log \int \mathrm{d}\mathbf{X} f_{k-1}(\mathbf{X})\mathrm{d}\mathbf{X}_\ell f_{k_\ell-1}(\mathbf{X}_\ell) \prod_{p=1}^{\ell-1}\mathrm{d}\mathbf{X}_p f_{k_p-2}(\mathbf{X}_p) \\ &\times \prod_{\mu} \left(Z(\boldsymbol{C}^*_{\ell+1,\mu})\right)^{n_\mu} \overline{\left(Z(\mu_\varepsilon|\boldsymbol{C}^*_{\ell+1,\mu})\right)}^{m_\mu} \\ &- \frac{1}{\ell+1}\log\mathcal{Z}_k - \frac{1}{\ell+1}\sum_{p=1}^{\ell}\log\mathcal{Z}_{k_p-1}, \end{aligned} \tag{D.34}$$

with

$$\boldsymbol{C}^*_{\ell+1,\mu} = \delta_{p,p'-1} + \delta_{p,p'+1} + (\mu_\varepsilon + x_p(\mu))\delta_{p,p'}. \tag{D.35}$$

In this case $\boldsymbol{C}^*_{\ell+1,\mu}$ corresponds to an embedded open chain of length ℓ (with $\ell+1$ nodes), with $k-2$ complex fields in the inner nodes, and $k-1$ fields on the first and last nodes.

For the case pure $\mathbf{B}$ traces we have the next expression,

$$\begin{aligned}\mathrm{Tr}(\mathbf{B}^{\ell}) &= \int \prod_{p=1}^{\ell} \mathrm{d}\boldsymbol{\varphi}_p \cdot \left[\prod_{p=1}^{\ell}\sum_{k_p} \tilde{p}(k_p)k_p\right] \prod_{p=1}^{\ell} \frac{1}{\mathcal{Z}_{k_p}\mathcal{Z}_{k_p-1}} \\ &\times \int \prod_{p=1}^{\ell} \mathrm{d}\boldsymbol{\psi}_p\mathrm{d}\mathbf{X}_p f_{k_p-1}(\mathbf{X}_p)\mathrm{d}\boldsymbol{Y}_p f_{k_p-1}(\boldsymbol{Y}_p)\, \mathrm{e}^{-\frac{\mathrm{i}}{2}\sum_p(\boldsymbol{\varphi}_p\cdot\mathbf{X}_p\boldsymbol{\varphi}_p - \frac{\mathrm{i}}{2}\boldsymbol{\psi}_{p-1}\cdot\boldsymbol{Y}_{p-1}\boldsymbol{\psi}_{p-1} - \mathrm{i}\boldsymbol{\psi}_{p-1}\cdot\boldsymbol{\varphi}_p)} \\ &= \left[\prod_{p=1}^{\ell}\sum_{k_p} \tilde{p}(k_p)k_p\right] \prod_{p=1}^{\ell} \mathrm{e}^{2C(k_{p-1},k_p)} \end{aligned} \tag{D.36}$$

We can keep up generalizing the mixed products,

$$\begin{aligned}
\mathrm{Tr}(\mathbf{M}^{\ell-1}\mathbf{B}^{m+1}) &= \int \prod_{p=1}^{\ell+m} \mathrm{d}\boldsymbol{\varphi}_p \left[\prod_{p=1}^{\ell+m}\sum_{k_p}\tilde{p}(k_p)\right] \prod_{p=1}^{\ell-1}\frac{k_p-1}{\mathcal{Z}_{k_p}} \prod_{p=\ell}^{\ell+m}\frac{k_p}{\mathcal{Z}_{k_p}\mathcal{Z}_{k_p-1}} \\
&\quad\times \int \prod_{p=1}^{\ell-1}\mathrm{d}\mathbf{X}_p f_{k_p-2}(\mathbf{X}_p) \prod_{p=\ell}^{\ell+m}\mathrm{d}\boldsymbol{\psi}_p \mathrm{d}\mathbf{X}_p f_{k_p-1}(\mathbf{X}_p)\mathrm{d}\boldsymbol{Y}_p f_{k_p-1}(\boldsymbol{Y}_p) \\
&\quad\times \mathrm{e}^{-\frac{\mathrm{i}}{2}\sum_{p=1}^{\ell-1}\boldsymbol{\varphi}_p\cdot\mathbf{X}_p\boldsymbol{\varphi}_p - \mathrm{i}\sum_{p=1}^{\ell-1}\boldsymbol{\varphi}_p\cdot\boldsymbol{\varphi}_{p+1} + \sum_{p=\ell}^{\ell+m}(-\frac{\mathrm{i}}{2}\boldsymbol{\varphi}_p\cdot\mathbf{X}_p\boldsymbol{\varphi}_p - \mathrm{i}\boldsymbol{\varphi}_{p+1}\cdot\boldsymbol{\psi}_p - \frac{\mathrm{i}}{2}\boldsymbol{\psi}_p\cdot\boldsymbol{Y}_p\boldsymbol{\psi}_p)} \\
&= \int \prod_{p=1}^{\ell+m}\mathrm{d}\boldsymbol{\varphi}_p \left[\prod_{p=1}^{\ell+m}\sum_{k_p}\tilde{p}(k_p)\right] \prod_{p=1}^{\ell-1}(k_p-1) \prod_{p=\ell}^{\ell+m}k_p \frac{1}{\mathcal{Z}_{k_{\ell+m}}} \prod_{p=1}^{\ell}\frac{1}{\mathcal{Z}_{k_p-1}} \prod_{p=\ell+1}^{\ell+m}\frac{1}{\mathcal{Z}_{k_{(p-1)}}\mathcal{Z}_{k_p-1}} \\
&\quad\times \int \prod_{p=1}^{\ell-1}\mathrm{d}\mathbf{X}_p f_{k_p-2}(\mathbf{X}_p) \prod_{p=\ell}^{\ell+m}\mathrm{d}\boldsymbol{\psi}_p \mathrm{d}\mathbf{X}_p f_{k_p-1}(\mathbf{X}_p)\mathrm{d}\boldsymbol{Y}_p f_{k_p-1}(\boldsymbol{Y}_p) \\
&\quad\times \mathrm{e}^{-\frac{\mathrm{i}}{2}\boldsymbol{\psi}_{\ell+m}\cdot\boldsymbol{Y}_{\ell+m}\boldsymbol{\psi}_{\ell+m} - \mathrm{i}\boldsymbol{\varphi}_1\cdot\boldsymbol{\psi}_{\ell+m} - \frac{\mathrm{i}}{2}\sum_{p=1}^{\ell}\boldsymbol{\varphi}_p\cdot\mathbf{X}_p\boldsymbol{\varphi}_p - \mathrm{i}\sum_{p=1}^{\ell-1}\boldsymbol{\varphi}_p\cdot\boldsymbol{\varphi}_{p+1}} \\
&\quad\times \mathrm{e}^{\sum_{p=\ell+1}^{\ell+m}(-\frac{\mathrm{i}}{2}\boldsymbol{\varphi}_p\cdot\mathbf{X}_p\boldsymbol{\varphi}_p - \mathrm{i}\boldsymbol{\varphi}_p\cdot\boldsymbol{\psi}_{p-1} - \frac{\mathrm{i}}{2}\boldsymbol{\psi}_{p-1}\cdot\boldsymbol{Y}_{p-1}\boldsymbol{\psi}_{p-1})} \\
&= \left[\prod_{p=1}^{\ell+m}\sum_{k_p}\tilde{p}(k_p)\right] \prod_{p=1}^{\ell-1}(k_p-1) \prod_{p=\ell}^{\ell+m}k_p \cdot \mathrm{e}^{(\ell+1)C(k_{\ell+m},\{k_p\}_{p=1,\dots,\ell}) + \sum_{p=\ell+1}^{\ell+m}2C(k_{p-1},k_p)}
\end{aligned} \tag{D.37}$$

So we can see the clear generalization, when $\ell = \sum_{t=1}^{s}\ell_t + m_t$

$$\begin{aligned}
\mathrm{Tr}(\prod_{t=1}^{s}\mathbf{M}^{\ell_t-1}\mathbf{B}^{m_t+1}) &= \left[\prod_{t=1}^{s}\prod_{p_t=1}^{\ell_t+m_t}\sum_{k_{p_t}}\tilde{p}(k_{p_t})\right] \prod_{t=1}^{s}\left[\prod_{p_t=1}^{\ell_t-1}(k_{p_t}-1)\prod_{p_t=\ell_t}^{\ell_t+m_t}k_{p_t}\right] \\
&\quad\times \exp\left(\sum_{t=1}^{s}[(\ell_t+1)C(k_{\ell_{t-1}+m_{t-1}},\{k_{p_t}\}_{p_t=1,\dots,\ell_t}) + \sum_{p_t=\ell_t+1}^{\ell_t+m_t}2C(k_{p_t-1},k_{p_t})]\right)
\end{aligned} \tag{D.38}$$

D.3.1 Imaginary replica limit of leading order terms

We now proceed to show the value of the RS expression of the previous sections once both the imaginary replica limit and the $\Delta \to 0$ limit are taken. Due to the RS we get expressions where we can take the limits according to (4.46) of Chapter 4. These expressions are of the form

$$\lim_{\Delta\to 0}\lim_{\substack{n_\mu\to \mathrm{i}\frac{\Delta}{\pi}\hat{\varrho}'(\mu)\\ m_\mu\to -n_\mu}}\prod_{\mu} f(\mu_\varepsilon)^{n_\mu}\overline{f(\mu_\varepsilon)}^{m_\mu} = \exp\left(\frac{2}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log f(\mu_\varepsilon)\right). \tag{D.39}$$

We note the following replica limits ,

$$\lim \log Z(\mathbf{X}) = -\frac{1}{\pi}\int \mathrm{d}\mu\, \hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu). \tag{D.40}$$

The saddle point equations become,

$$\begin{aligned} W[\boldsymbol{x}] =& \frac{1}{\mathcal{C}^2}\sum_k p(k)\frac{k}{c}\frac{1}{\mathcal{Z}_k} f_{k-1}[\boldsymbol{x}]\mathrm{e}^{-\frac{1}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu)} \\ \mathcal{C}^2 =& \sum_k p(k)\frac{k}{c}\frac{\mathcal{Z}_{k-1}}{\mathcal{Z}_k} \\ \mathcal{Z}_k =& \int \mathcal{D}\boldsymbol{x}\, f_k[\boldsymbol{x}]\mathrm{e}^{-\frac{1}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu)}, \end{aligned} \tag{D.41}$$

and the action

$$\begin{aligned} S_0[W] =& -\frac{c}{2}\log\int \mathcal{D}\boldsymbol{x}\mathcal{D}\boldsymbol{x}'\, W[\boldsymbol{x}]W[\boldsymbol{x}']\mathrm{e}^{-\frac{1}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log(1-1/(x(\mu)x'(\mu)))} \\ &+\sum_k p(k)\log\int \mathcal{D}\boldsymbol{x}\, f_k[\boldsymbol{x}]\mathrm{e}^{\cdot} \end{aligned} \tag{D.42}$$

Note that we can define and effective degree distribution for the equations and rewrite the equations in an even more compact form,

$$\begin{aligned} \tilde{p}(k) =& \frac{1}{\mathcal{C}^2}p(k)\frac{k}{c}\frac{\mathcal{Z}_{k-1}}{\mathcal{Z}_k} \\ W[\boldsymbol{x}] =& \sum_k \tilde{p}(k)\frac{f_{k-1}[\boldsymbol{x}]A[\boldsymbol{x}]}{\mathcal{Z}_{k-1}}. \end{aligned} \tag{D.43}$$

It is in this form that the weighted population dynamics algorithm presented in (6.2.1) can be clearly understood.

D.3.2 Imaginary replica limits of subleading order terms

For the RS form of the subleading order terms it is only necessary to calculate the replica limit of function $R(\{k_p\}_{p=1,\ldots,\ell})$ and $C(k,\{k_p\}_{p=1,\ldots,\ell})$ since we proved in the previous section that all the different traces can be expressed as combinations of these two functions.

The result is,

$$\begin{aligned}
\lim R(\{k_p\}_{p=1,\ldots,\ell}) =& \frac{1}{\ell}\log\int\prod_{p=1}^{\ell}\mathcal{D}\boldsymbol{x}_p f_{k_p-2}[\boldsymbol{x}_p]\mathrm{e}^{\ell\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}g_\ell(\mu)} \\
&- \frac{1}{\ell}\sum_{p=1}^{\ell}\log\int\mathcal{D}\boldsymbol{x} f_{k_p-1}[\boldsymbol{x}]\mathrm{e}^{-\frac{1}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu)} \\
\lim C(k,\{k_p\}_{p=1,\ldots,\ell}) =& \frac{1}{\ell+1}\log\int\mathcal{D}\boldsymbol{x} f_{k-1}[\boldsymbol{x}]\mathcal{D}\boldsymbol{x}_\ell f_{k_\ell-1}[\boldsymbol{x}_\ell] \\
&\times\prod_{p=1}^{\ell-1}\mathcal{D}\boldsymbol{x}_p f_{k_p-2}[\boldsymbol{x}_p]\mathrm{e}^{\ell\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}c_\ell(\mu)} \\
&- \frac{1}{\ell+1}\log\int\mathcal{D}\boldsymbol{x} f_k[\boldsymbol{x}]\mathrm{e}^{-\frac{1}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu)} \\
&- \frac{1}{\ell+1}\sum_{p=1}^{\ell}\log\int\mathcal{D}\boldsymbol{x} f_{k_p-1}[\boldsymbol{x}]\mathrm{e}^{-\frac{1}{\pi}\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu)}
\end{aligned} \tag{D.44}$$

Where we have defined

$$\begin{aligned}
g_\ell(\mu) =& g_\ell(\mu|\{x_p\}_{p=1,\ldots,\ell}) = \frac{2}{\pi\ell}\mathrm{Im}\log Z(\mu|\mathbf{A}^*_{\ell,\mu}) \\
c_\ell(\mu) =& c_\ell(\mu|x,\{x_p\}_{p=1,\ldots,\ell}) = \frac{2}{\pi(\ell+1)}\mathrm{Im}\log Z(\mu|\boldsymbol{C}^*_{\ell+1,\mu})
\end{aligned} \tag{D.45}$$

D.3.3 Derivatives with respect to $\hat{\varrho}(\mu)$ in the $\delta W/\delta\hat{\varrho}(\mu)=0$ direction.

For the pure $\mathbf{M}$ traces,

$$\begin{aligned}
\frac{\delta}{\delta\hat{\varrho}(\mu)}\lim\mathrm{Tr}(\mathbf{M}^\ell) =& \left[\prod_{p=1}^{\ell}\sum_{k_p}\tilde{p}(k_p)(k_p-1)\right]\exp\left(\ell R(\{k_p\}_{p=1,\ldots,\ell})\right) \\
&\times\left[\left\langle\ell\frac{\mathrm{d}}{\mathrm{d}\mu}g_\ell(\mu|\{x_p\}_{p=1,\ldots,\ell})\right\rangle_{W_R} - \sum_{p=1}^{\ell}\left\langle\frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu)\right\rangle\right]
\end{aligned} \tag{D.46}$$

Where we have introduced,

$$W_R[\{(k_p,\boldsymbol{x}_p)\}_{p=1,\ldots,\ell}] = \frac{\prod_{p=1}^{\ell}f_{k_p-2}[\boldsymbol{x}_p]\ \cdot\ \mathrm{e}^{\ell\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}g_\ell(\mu)}}{\int\prod_{p=1}^{\ell}\mathcal{D}\boldsymbol{x}_p f_{k_p-2}[\boldsymbol{x}_p]\mathrm{e}^{\ell\int \mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}g_\ell(\mu)}}. \tag{D.47}$$

For the mixed case,

$$\frac{\delta}{\delta\hat{\varrho}(\mu)}\lim \mathrm{Tr}(\prod_{t=1}^{s}\mathbf{M}^{\ell_t-1}\mathbf{B}^{m_t+1}) = \left[\prod_{t=1}^{s}\prod_{p_t=1}^{\ell_t+m_t}\sum_{k_{p_t}}\tilde{p}(k_{p_t})\right]\prod_{t=1}^{s}\left[\prod_{p_t=1}^{\ell_t-1}(k_{p_t}-1)\prod_{p_t=\ell_t}^{\ell_t+m_t}k_{p_t}\right]$$
$$\times\exp\left(\sum_{t=1}^{s}\left[(\ell_t+1)C(k_{\ell_{t-1}+m_{t-1}},\{k_{p_t}\}_{p_t=1,\ldots,\ell_t})+\sum_{p_t=\ell_t+1}^{\ell_t+m_t}2C(k_{p_t-1},k_{p_t})\right]\right)$$
$$\times\left[\sum_{t=1}^{s}\left(\left\langle(\ell_t+1)\frac{\mathrm{d}}{\mathrm{d}\mu}c_\ell(\mu|x,\{x_p\}_{p=1,\ldots,\ell})\right\rangle_{W_C}\right.\right.$$
$$-\left\langle\frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu)\right\rangle_{W_{k_{\ell_{t-1}+m_{t-1}}}}-\sum_{p_t=1}^{\ell_t}\left\langle\frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu)\right\rangle_{W_{k_{p_t}-1}}$$
$$\left.\left.+\sum_{p_t=\ell_t+1}^{\ell_t+m_t}\left\langle 2\frac{\mathrm{d}}{\mathrm{d}\mu}c_1(\mu|x,x_1)\right\rangle_{W_C}-\left\langle\frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu)\right\rangle_{k_{(p_t-1)}}-\left\langle\frac{\mathrm{d}}{\mathrm{d}\mu}h(\mu)\right\rangle_{k_{p_t}-1}\right)\right] \tag{D.48}$$

With,

$$W_C[(k,\boldsymbol{x}),\{(k_p,\boldsymbol{x}_p)\}_{p=1,\ldots,\ell}] = \frac{f_{k-1}[\boldsymbol{x}]f_{k_\ell-1}[\boldsymbol{x}_\ell]\prod_{p=1}^{\ell-1}f_{k_p-2}[\boldsymbol{x}_p]\,\mathrm{e}^{\ell\int\mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}c_\ell(\mu|x,\{x_p\}_{p=1,\ldots,\ell})}}{\mathcal{Z}_C}. \tag{D.49}$$

Where again,

$$\mathcal{Z}_C = \int\mathcal{D}\boldsymbol{x}f_{k-1}[\boldsymbol{x}]\mathcal{D}\boldsymbol{x}_\ell f_{k_\ell-1}[\boldsymbol{x}_\ell]\prod_{p=1}^{\ell-1}\mathcal{D}\boldsymbol{x}_p f_{k_p-2}[\boldsymbol{x}_p]\mathrm{e}^{\ell\int\mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}c_\ell(\mu|x,\{x_p\}_{p=1,\ldots,\ell})}. \tag{D.50}$$

Therefore, the generating function in the replica limit is given by the following formula,

$$\phi[\hat{\varrho}] = -\frac{c}{2}\log\int\mathcal{D}\boldsymbol{x}\mathcal{D}\boldsymbol{x}'W[\boldsymbol{x}]W[\boldsymbol{x}']\mathrm{e}^{-\frac{1}{\pi}\int\mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log(1-\frac{1}{x(\mu)x'(\mu)})}$$
$$+\sum_k p(k)\log\int\mathcal{D}\boldsymbol{x}f_k[\boldsymbol{x}]\mathrm{e}^{-\frac{1}{\pi}\int\mathrm{d}\mu\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu)}$$
$$+\lim\frac{1}{N}\left(\frac{c^2}{4}-\frac{1}{2}\mathrm{Tr}(\mathbf{B})-\mathrm{Tr}(\mathbf{MB})+\frac{1}{4}\mathrm{Tr}(\mathbf{B}^2)+\sum_{\ell=3}^{\infty}\frac{1}{2\ell}\mathrm{Tr}((\mathbf{M}-\mathbf{B})^\ell)\right) \tag{D.51}$$

and provided $\delta W/\delta\hat{\varrho}(\mu) = 0$, the spectrum formula is

$$\begin{aligned}\varrho(\mu) = \frac{\delta\phi[\hat{\varrho}]}{\delta\hat{\varrho}(\mu)} = -\frac{c}{2}\int \mathcal{D}\boldsymbol{x}\mathcal{D}\boldsymbol{x}' W_{\mathcal{K}}[\boldsymbol{x},\boldsymbol{x}']\left(-\frac{1}{\pi}\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log(1-1/(x(\mu)x'(\mu)))\right)\\ + \sum_k p(k)\int \mathcal{D}\boldsymbol{x} W_k[\boldsymbol{x}]\left(-\frac{1}{\pi}\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{Im}\log x(\mu)\right)\\ + \frac{1}{N}\frac{\delta}{\delta\hat{\varrho}(\mu)}\lim\left[-\frac{1}{2}\mathrm{Tr}(\mathbf{B}) - \frac{1}{2}\mathrm{Tr}(\mathbf{MB}) + \frac{1}{4}\mathrm{Tr}(\mathbf{B}^2) + \sum_{\ell=3}^{\infty}\frac{1}{2\ell}\mathrm{Tr}((\mathbf{M}-\mathbf{B})^{\ell})\right].\end{aligned} \tag{D.52}$$

Where one only needs to expand the traces and substitute (D.46) and (D.48) accordingly.